An Investigation of the DNA Polymerases Responsible for Translesion Synthesis and DNA Repair

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for Graduation with *Research Distinction in Biochemistry* in the College of Arts and Sciences of the Ohio State University

Lindsey Renee Pack

The Ohio State University

June 2009

Project Advisor: Dr. Zucai Suo, Department of Biochemistry

Contents

Acknowledgements

I would like to take the opportunity to thank those who have helped me to achieve the work presented in this senior honors thesis. I would first like to thank Dr. Zucai Suo for seeing my potential and inviting me to join his laboratory. He has continued to support my research and the pursuit of my goals, especially my desire to continue my education in graduate school. Next, I would like to acknowledge and thank my lab mates, which include Jessica Brown, Shanen Sherrer, Jason Fowler, Sean Newminster, John Pryor, Kevin Fiala, Amy Xu, Leonardo Porchia, David Beyer, and Laura Sanman. I would like to especially acknowledge Shanen Sherrer and Jessica Brown for allowing me to take part in many of their research projects. They both have offered great mentoring and friendship. I am grateful to them for the research training that they have provided to me and for their continued support in my research endeavors.

Outside of my laboratory group, I would like to thank two of my mentors, Dr. Jennifer Ottesen and Dr. Venkat Gopalan. Both of these outstanding individuals were instrumental in my decision to attend graduate school and the success I had in applying for admission to various universities and fellowships. I will not forget their continued support and encouragement throughout my undergraduate career.

Finally, I would like to thank my family and friends. First, I would like to thank my parents for always encouraging me and reminding me how proud they are of me. My parents raised me to be the best I can be. My hard work and success as an undergraduate are a reflection of the values that my parents have instilled in me. I would also like to thank my brothers, Aaron and Ryan for being amazing individuals and for making me so proud to be their sister. I must also thank Aaron for bringing me food when I was hard at work in the lab. In addition to my

family I would like to thank my friends, Alysa Kociuruba, Elizabeth Johnson, Jean Wheasler, Anna Bryant, and Mridula Manohar. These amazing women have helped me to grow as a person throughout my college career. They have provided invaluable encouragement and support, especially during the times when my lab work was frustrating and exhausting. Despite the end of our undergraduate education, I know that we will remain friends for life. Finally, I would like to thank my boyfriend John Hanson for being the person who I could tell everything to. His support has made me believe that I can achieve great things. Finally, I want to thank him for calling me a lab champion.

Abstract

The replication of DNA is vital to the survival of any organism. The ability to replicate DNA, however, is compromised by the continual onslaught of DNA damaging agents, which include UV light, oxidative stress, and a variety of environmental pollutants. DNA damage stalls the replication of the genome by the normal replicative polymerases. Thus, in order for cell survival, this damage must either be repaired or overcome by translesion synthesis. The X- and Y- families of DNA polymerases have been implicated in these cell survival pathways. These families of DNA polymerases have roles in both the bypass of DNA lesions as well as filling gaps in DNA synthesis. Despite the significance of their function, many aspects of the mechanism by which these DNA polymerases overcome DNA damage are largely unknown. Thus, there is an obvious need to investigate the structure, function, and the mechanism of translesion synthesis of specific X- and Y- family polymerases.

The first part of this thesis will investigate how Y-family DNA polymerases are able to bypass a specific lesion, known as *N*-(deoxyguanosin-8-yl)-1-aminopyrene (dG^{AP}), in the DNA. The first chapter will report on the determination of the kinetic mechanism of translesion synthesis opposite this lesion by a model Y-family DNA polymerase, *Sulfolobus sulfataricus* DNA polymerase IV (Dpo4). Pre-steady-state kinetic methods were employed to investigate the incorporation of individual nucleotides upstream, downstream, and opposite the site of the dG^{AP} lesion. Our results showed that Dpo4 was able to bypass this bulky lesion, but was inefficient in its replication opposite the lesion and immediately downstream of the lesion. The inefficient bypass also corresponded to a decrease in enzyme fidelity at these two sites. Moreover, our studies indicated that the kinetics of nucleotide incorporation at these sites was biphasic, in

which a small, fast phase was followed by a larger, slow phase. The results from the various kinetic studies performed to investigate Dpo4's bypass of the dG^{AP} lesion are reported in the first chapter. These results have also been previously published in the Journal of Biological Chemistry [1].

The second part of the thesis describes the efficiency and mutagenic characterization of the bypass of the dG^{AP} lesion by the model Y-family DNA polymerase Dpo4, as well as two human Y-family DNA Polymerases, human DNA polymerases kappa and eta. To determine the mutagenic profile for each of these enzymes, a novel assay, known as Short Oligonucleotide Sequencing Assay (SOSA), was utilized. First, running start assays established that all three Y-family polymerases were capable of bypassing the dG^{AP} lesion. It was shown that DNA polymerase eta did so with the most efficiency. However, when comparing the mutagenic profiles of these three polymerases, it was shown that Dpo4 had the greatest fidelity. When comparing the human Y-family DNA polymerases, human DNA polymerase kappa had an overall greater fidelity than DNA polymerase eta. These results are discussed in Chapter 2.

In the third part of the thesis, we determined if nucleoside analogs, a class of antiviral drug, are potential substrates for human X- and Y- family polymerases. Human DNA polymerase λ, which putatively functions in base excision repair, and human DNA polymerase η, which functions in translesion synthesis, were chosen as representative members of the X- and Y- family enzymes, respectively. Several major drugs, including entecavir, zidovudine, tenofovir, adefovir, emtricitabine, and ribovarin, were chosen, since they are currently being used in hepatitis B and HIV-1 clinical virology to treat the millions of people infected worldwide. Single nucleotide incorporation assays were used to measure the substrate specificity

for the natural nucleotides and their respective nucleotide analog. By comparing these substrate specificities, the probability of the nucleotide analog being incorporated into DNA by DNA polymerases λ and η could be determined. Interestingly, the order of substrate preference was different for Polλ (emtricitabine< tenofovir< adefovir<zidovudine< entecavir) and Polη (tenofovir<zidovudine<emtricitabine<adefovir<entecavir). Kinetic results suggest that X- and Y-family DNA polymerases may incorporate these drugs *in vivo*.

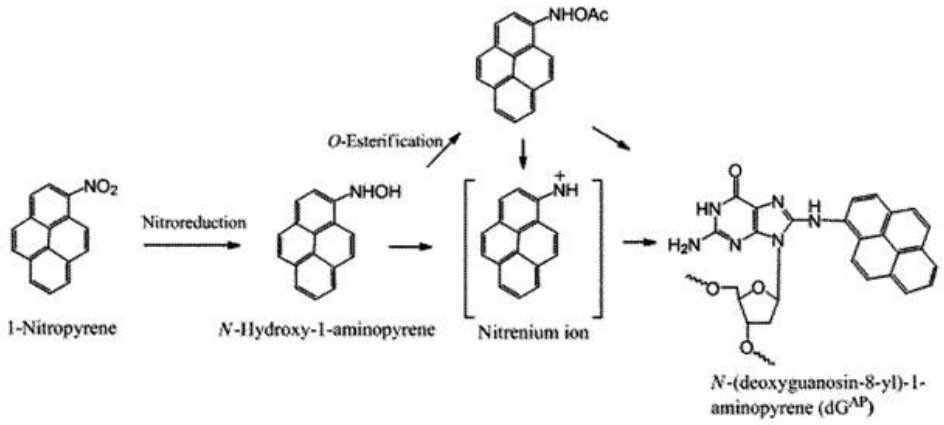
Chapter 1: Bypass Mechanism of a Lesion Formed from 1-Nitropyrene by a Y-family DNA Polymerase

Background and Introduction

As previously stated, replication of genomic material is essential to the growth and reproduction of any organism. The process of DNA replication is complicated and compromised by DNA damaging agents, which include UV light, oxidative stress, and a variety of environmental pollutants [2, 3]. Environmental pollutants have been shown to affect human health at the molecular level. For example, they can modify genomic DNA by introducing lesions to nucleotides [3]. These lesions will stall the principal replicative DNA polymerases and thus inhibit the replication of the cellular genome [4, 5]. The cell has two basic ways to overcome damaged DNA and the stalling of replication: i) Recognition of the lesion and removal by the cellular DNA repair machinery or ii) Displacement of the principal replicative DNA polymerase and continued polymerization by lesion bypass DNA polymerases, which transverse the unrepaired lesions. The Y-family DNA polymerases are principally responsible for translesion synthesis due to their relatively flexible and solvent accessible active sites, though, the Y-family DNA polymerases synthesize DNA with a low fidelity and processivity [6]. The Y-family of DNA polymerases are found in the three domains of life - Archea, Bacteria, and Eukaryota. In humans, four Y-family DNA polymerases have been identified (DNA polymerases Rev1, η, ι, and κ). *Sulfolobus solfataricus* DNA polymerase IV (Dpo4), a model Y-family DNA polymerase was utilized for the following study. Dpo4 was chosen for its high thermostability, relative ease of expression in *E. coli* and subsequent purification, as well as

because it has been extensively studied *in vitro* [7]. In addition, Dpo4 has been shown to be capable of bypassing a variety of DNA lesions including apurinic/apyrimidinic (abasic) sites [8], cis-syn thymine-thymine dimers [9], cisplatin-induced 1,2-intrastrand cross links with adjacent deoxyguanosines [10], and benzo[a]pyrene diol epoxides on deoxyguanosines (BPDE-dG) [11].

Dpo4 will be utilized to study the bypass of a 1-nitropyrene (1-NP)-induced lesion because there have been no mechanistic studies of this process *in vitro*. 1-NP, an abundant polycyclic aromatic hydrocarbon (PAH) compound, is produced due to incomplete diesel and gasoline combustion [2, 3]. Metabolism of 1-NP takes place through two known pathways: nitroreduction (Scheme 1.1) and C-hydroxylation. If 1-NP is oxidized into non-DNA-reactive metabolites by P40 enzymes while in the gastrointestinal tract, respiratory system or skin, the organism can excrete the metabolites through a detoxification process [12, 13]. However, if the 1-NP proceeds through nitro reduction, DNA-reactive metabolites will be formed (Scheme 1.1) [12]. As seen in Scheme 1.1, the major product formed from the DNA-reactive metabolites is a DNA lesion known as *N*-(deoxyguanosin-8-yl)-1-aminopyrene (dG^{AP}). This lesion has been shown to be mutagenic in bacterial and mammalian cells [12, 14]. The International Agency for Research on Cancer has classified 1-NP as a class 2B carcinogen [12, 14].



Scheme 1.1: Metabolism of 1-nitropryene via the nitroreduction pathway

Dpo4 was chosen to study the bypass of dG^{AP} for a variety of reasons, including: i) the kinetic mechanism of nucleotide incorporation into undamaged DNA was previously elucidated in the Suo laboratory [15], ii) many crystal structures of Dpo4 in ternary complexes that contain various lesions have been published [16-18], iii) Dpo4 is the only Y-family DNA polymerase encoded in the *Sulfolobus sulfataricus* genome, and thus would be responsible for most of its translesion DNA synthesis [19], and finally and most importantly iv) Dpo4 is a good model for eukaryotic Y-family DNA polymerases, like human DNA polymerase κ [20]. Studying Dpo4's kinetic mechanism of bypass could provide insight into how eukaryotic Y-family DNA polymerases bypass dG^{AP}. To understand the bypass mechanism of a Y-family DNA polymerase catalyzing DNA synthesis opposite dG^{AP} and to investigate the potential mutagenicity of 1-NP, an oligonucleotide was synthesized placing the dG^{AP} single base lesion in a region rich in GCs. Regions composed of repetitive DNA sequence have been shown to induce more mutations than non-repetitive regions [14, 21]. To understand the mechanism of

bypass of the dG^{AP} lesion catalyzed by Dpo4, pre-steady-state kinetic methods were utilized.

Methods

Materials- For this project, OptiKinase from United Biochemical (Cleveland, OH), [γ-^{32}P]ATP from GE Healthcare, and nucleotides (dNTPs) from Invitrogen were purchased. The full length Dpo4 was expressed in *E. Coli* and purified within the Suo laboratory as previously described [7].

Synthetic Oligonucleotides-The DNA template 26-mer dG^{AP} was synthesized as previously described [22]. Other DNA templates and primers (Table 1.1) were purchased from Integrated DNA Technologies (Coralville, IA) and purified using denaturing PAGE. The concentration of these DNA oligonucleotides was determined using UV spectroscopy at 260 nm.

Table 1.1 Sequences of DNA oligonucleotides

Primers	
17-mer	5′-AACGACGGCCAGTGAAT-3′
19-mer	5′-AACGACGGCCAGTGAATTC-3′
20-mer	5′-AACGACGGCCAGTGAATTCG-3′
21-mer	5′-AACGACGGCCAGTGAATTCGC-3′
22-mer	5′-AACGACGGCCAGTGAATTCGCG-3′
23-mer	5′-AACGACGGCCAGTGAATTCGCGC-3′
Templates	
26-mer	3′-TTGCTGCCGGTCACTTAAGCGCGCCC-5′
[a]26-mer-dG^{AP}	3′-TTGCTGCCGGTCACTTAAGC**G**CGCCC-5′
DNA Trap	
D-1 (21/41-mer)	5′-CGCAGCCGTCCAACCAACTCA-3′ 3′-GCGTCGGCAGGTTGGTTGAGTAGCAGCTAGGTTACGGCAGG-5′

[a]**G** designates *N*-(deoxyguanosin-8-yl)-1-aminopyrene (dG^{AP}).

Radiolabeling and Annealing DNA substrates-Each primer was 5'-^{32}P-labeled by incubating it with Optikinase and [γ-^{32}P]ATP for 3 hours at 37 ˚C. These radiolabeled primers were annealed to either 26-mer or 26-mer-dG^{AP}, unlabeled templates, at a molar ratio of 1.00:1.15. This mixture was heated to 75 ˚C for two minutes and then slowly cooled to room temperature.

Buffers: All pre-steady-state kinetic assays were performed in optimized buffer R (50 mM HEPES, pH 7.5 at 37 ˚C, 5 mM $MgCl_2$, 50 mM NaCl, 0.1 mM EDTA, 5mM dithiothreitol, 10% glycerol, and 0.1 mg/mL bovine serum albumin) [7, 15]. Electrophoresis mobility shift assays (EMSA) were performed in buffer S (50 mM Tris-Cl, pH 7.5 at 23 ˚C, 5 mM $MgCl_2$, 50 mM NaCl. 5 mM dithiothreitol, 10% glycerol, and 0.1 mg/mL bovine serum albumin). The concentrations indicated above were those of the final mixed solutions.

Running Start Assay-The running start assay was performed as previously described [8, 10]by a Jessica Brown, a colleague in the laboratory. In brief, a solution of 5'-^{32}P-labeled DNA (100 nM) and Dpo4 (100 nM) in buffer R was pre-incubated at 37 ˚C and subsequently rapidly mixed with a solution containing all four dNTPs (200 μM each) at 37 ˚C via a rapid chemical-quench flow apparatus (KinTek). The reaction was quenched with .37 M EDTA after various reaction times. The reaction products were visualized by denaturing PAGE (17% polyacrylamide, 8 M urea).

EMSA- Dpo4 (0.5-80 nM) was titrated into a solution containing 5'-^{32}P-labeled DNA (5nM) in buffer S at 23 ˚C. To separate the free DNA from DNA complexed with Dpo4, native PAGE was performed at 70 V for 35 minutes at 23 ˚C using running buffer A (50 mM Tris acetate, pH 7.5, at 23 ˚C, 0.5 mM EDTA, 5.5 $Mg(OAc)_2$). After drying the gel, the bands were

quantitated using PhosphorImager 445 SI (Molecular Dynamics). The dependence of the concentration of complexed Dpo4 and DNA (Dpo4·DNA) on the concentration of Dpo4 was fit using Equation 1.1 to yield $K_{d,\ DNA}$, the equilibrium dissociation constant for the binary complex(Dpo4·DNA).

$$[\text{Dpo4}\cdot\text{DNA}] = 0.5(K_{d,DNA} + E_0 + D_0) - 0.5[K_{d,\ DNA} + E_0 + D_0)^2 - 4E_0D_0]^{1/2} \quad \text{(Eq 1.1)}$$

In Equation 1.1, E_0 is the active Dpo4 concentration and D_0 is the DNA concentration.

Determination of Substrate Specificity – Pre-steady state single turnover dNTP incorporation assays were employed to determine the k_p and $K_{d,\ dNTP}$, where the k_p is the maximum rate of incorporation of a dNTP and $K_{d,\ dNTP}$ is the equilibrium dissociation constant of an incoming nucleotide. These assays were performed as previously described [7, 15]. In brief, a solution of 5'-^{32}P-labeled DNA (30 nM) and Dpo4 (120 nM) in buffer R was pre-incubated at 37 ˚C and subsequently mixed with solutions containing increasing amounts of a single dNTP. These reactions were quenched using the rapid-chemical quench or manually after various time points using 0.37 M EDTA. The formed product was separated from the initial primer by denaturing PAGE (17% acrylamide, 8 M urea) and quantitated with a Phosphorimager 445 SI. For each concentration of dNTP, the function of product formation per time point was fit to a single exponential equation (Equation 1.2).

$$[\text{Product}] = A(1-\exp(-k_{obs}*t)) \quad \text{(Eq 1.2)}$$

Where k_{obs} is the observed reaction rate constant, and A is the reaction's product formation amplitude. Next, the values of the k_{obs} were plotted against the concentration of dNTP for the respective reaction. This relationship was fit to a hyperbolic equation (Equation 1.3).

$$k_{obs} = k_p[\text{dNTP}]/\{[\text{dNTP}] + K_{d,\,DNA}\} \quad \text{(Eq 1.3)}$$

Where, as stated previously, k_p is the maximum rate of dNTP incorporation, and the $K_{d,dNTP}$ is the dissociation constant for the ternary complex (Dpo4·DNA·dNTP) at equilibrium. The substrate specificity is the ratio of the k_p over the $K_{d,dNTP}$.

Biphasic Kinetic Assay- Shanen Sherrer, a colleague in the Suo laboratory, performed the biphasic kinetic assay. In short, a pre-incubated solution of Dpo4 (120 nM) and 5'-^{32}P-labeled DNA (30 nM) in buffer R was rapidly mixed with 5 μM DNA trap D-1 (Table 1.1) [7] and 1.2 mM correct dNTP in buffer R. After various time points, the mixture was quenched with 0.37 M EDTA. The reaction products were separated using PAGE and visualized using the Phosphorimager 445 SI. The product formation versus reaction time was fit using a double-exponential equation (Equation 1.4):

$$[\text{Product}] = E_0A_1[1-\exp(-k_1t)] + E_0A_2[1-\exp(-k_2t)] \quad \text{(Eq 1.4)}$$

Where E_0 is the active concentration of Dpo4, A_1 and A_2 are the reaction amplitudes of the two phases, and k_1 and k_2 are the rate constants for the two phases.

Results

Running Start Assay- A running start assay was performed to observe the DNA polymerization pattern of Dpo4 in the presence of a dG^{AP} lesion in the DNA substrate (17/26-mer-dG^{AP}) and to compare the polymerization with a controlled DNA substrate (17/26-mer). As demonstrated previously, Dpo4 synthesized full-length product 26-mer with the undamaged DNA substrate 17/26-mer (Table 1.1) within 10 s (Figure 1.1A) [10]. While, Dpo4 was capable of synthesis bypassing the dG^{AP} lesion, the enzyme was significantly slowed and full-length

product was not observed until after 180 s (Figure 1.1B). Interestingly, during synthesis, an accumulation of intermediate products, 20- and 21-mer, indicated that Dpo4 stalled in its dNTP incorporation directly opposite the dG^{AP} lesion and in the extension of the lesion bypass product, respectively. No pausing was demonstrated in the DNA synthesis opposite the control template. In Figure 1.1A, formation of the 27-mer was likely due to blunt end addition [23], while accumulation of 24 and 25-mer in Figure 1.1A and 25-mer in Figure 1.1B was likely caused by polymerase "slippage" via primer realignment.

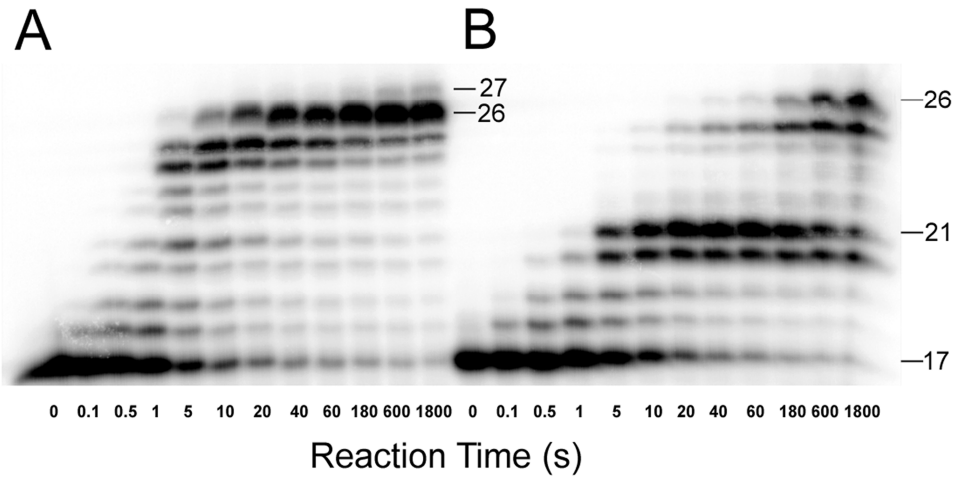


Figure 1.1: Running start assay for Dpo4. (A) 17/26-mer; (B) 17/26-mer-dG^{AP}. Sizes of important products are indicated, and the 21st position marks the location of the dG^{AP} lesion from the 3'-terminus of the DNA template.

DNA Binding of Dpo4 Effected by dG^{AP}-The accumulation of product at the pause sites in Figure 1.1B suggested that the presence of dG^{AP}, a bulky lesion, may cause a weakening in the binding affinity of Dpo4 to the DNA. To measure the binding affinity ($1/K_{d,DNA}$) of DNA to Dpo4, EMSA was performed for both control and damaged DNA templates. PAGE was utilized to separate the binary complex (DNA·Dpo4) from unbound DNA (Figure 1.2).

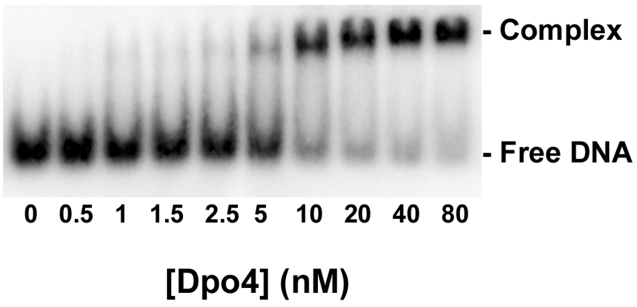


Figure 1.2: EMSA gel image. The binary complex of Dpo4•DNA was separated from free DNA by native PAGE.

The concentration of the Dpo4·DNA complex was plotted against the total concentration of Dpo4 and fit to Equation 1.1. This quantitation is exemplified in Figure 1.3, where the DNA substrate was the 20/26-mer-dG^{AP} and the $K_{d,DNA}$ was determined to be 1.0 ± 0.1 nM. An EMSA was used to determine the binding affinity of DNA and Dpo4 for the controlled and damaged DNA substrates as listed in Table 1.2. As expected, Dpo4 bound to the different substrate sites of undamaged DNA with approximately the same binding affinity (3.1-4.0 nM). In the presence of the damaged DNA substrates, a larger range of $K_{d,DNA}$ values was determined (1.0-4.4 nM). At the first pause site (20/26-mer-dG^{AP}) a four-fold tighter binding was observed. This tightening of the Dpo4·DNA complex could suggest that the dG^{AP} moiety may interact directly with the active site residues of Dpo4. Typically, though, an increase in binding affinity should facilitate processive polymerization and, thus, the tight binding of 20/26-mer-dG^{AP} cannot account for the accumulation of 20-mer. To further investigate the accumulation of intermediates in the vicinity of a DNA lesion, the microscopic kinetic parameters of dNTP incorporation, the maximum rate of incorporation (k_p) and the dNTP binding affinity ($1/K_{d,DNA}$), were studied.

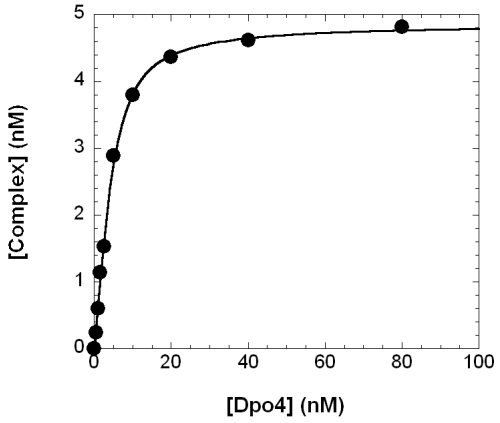


Figure 1.3: EMSA plot. The binary complex's concentration versus the total concentration of Dpo4. The data was fit to Eq 1.1.

Table 1.2 Binding affinity of Dpo4 to damaged and control DNA substrates at 23 °C

DNA Substrate	Damaged DNA[a] (nM)	Control DNA[b] (nM)	Affinity Ratio[c]
19/26-mer	2.7 ± 0.2	3.1 ± 0.5	1.2
20/26-mer	1.0 ± 0.1	4.0 ± 0.2	4.0
21/26-mer	4.4 ± 0.2	3.7 ± 0.2	0.8
22/26-mer	2.6 ± 0.2	3.8 ± 0.6	1.5
23/26-mer	4.0 ± 0.3	3.6 ± 0.5	0.9

[a]Damaged DNA refers to those with template 26-mer-dG^{AP} in Table 1.
[b]Control DNA refers to those with template 26-mer in Table 1.
[c]Calculated as $(K_{d,\,DNA})_{Control}/(K_{d,\,DNA})_{Damaged}$.

The Kinetics of a dG^{AP} lesion and its Effect on dNTP Incorporation- In order to determine the k_p and $K_{d,DNA}$, we performed single dNTP incorporation assays under single-turnover reaction conditions. In this assay, a pre-incubated solution of Dpo4 (120 nM) and 5'-^{32}P-labeled DNA substrate are rapidly mixed with a dNTP and quenched at various time points with .37M EDTA. The products are then separated using PAGE. The product formation is then

plotted against time and fit to Equation 1.2 to determine the observed reaction rate (k_{obs}) (Figure 1.4). The dependence of k_{obs} on the concentration of the dNTP is then plotted and fit to Equation 1.3, which yields values for the k_p and $K_{d,\,DNA}$ (Figure 1.5).

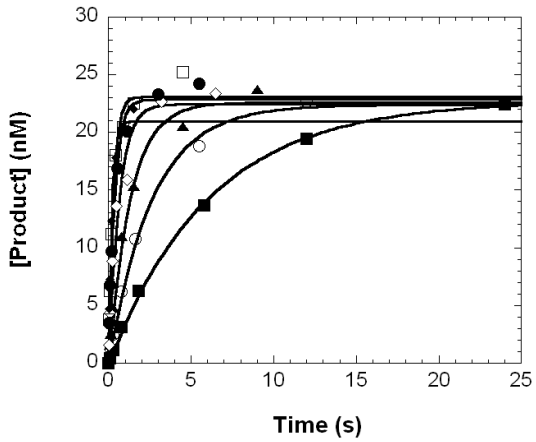


Figure 1.4: Relationship of Product Formation versus Time. Each time course was fit to Eq 1.2 to yield a k_{obs}.

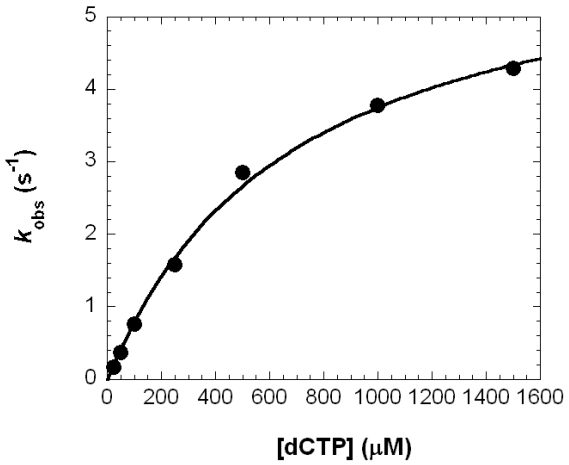


Figure 1.5: The Plot of Observed Rate Constant against Concentration. The plot of k_{obs} values against dCTP concentrations was fit to Eq 1.3 to produce a k_p of 6.3 ± 0.3 s^{-1} and a $K_{d,\,dCTP}$ of 682 ± 80 μM.

This single dNTP incorporation assay was performed for the series of DNA substrates representing the progression of Dpo4 as it approached, encountered, and bypassed the dG^{AP} lesion. This data is found in Table 1.3. Besides performing these assays on the damaged substrates, we also performed them on the control template 26-mer in order to generate a comparison (Appendix 1).

Table 1.3 Kinetic parameters for single dNTP incorporation opposite template 26-mer-dG^{AP}

dNTP	$K_{d,\ dNTP}$ (μM)	k_p (s^{-1})	$(k_p/K_{d,\ dNTP})_{damaged}$ ($\mu M^{-1}s^{-1}$)	Efficiency Ratio[a,b]	Fidelity[c]	Probability[d]
19/26-mer-dG^{AP}						
dGTP	187 ± 36	4.3 ± 0.3	$2.3x10^{-2}$	1.1	-	99.8
dATP	373 ± 81	$(6.3 ± 0.5)x10^{-3}$	$1.7x10^{-5}$	0.8	$7.4x10^{-5}$	0.1
dCTP	227 ± 23	$(6.4 ± 0.2)x10^{-3}$	$2.8x10^{-5}$	5	$1.2x10^{-3}$	0.1
dTTP	1180 ± 190	$(1.0 ± 0.1)x10^{-2}$	$8.7x10^{-6}$	1.0	$3.8x10^{-4}$	0.0
*20/26-mer-dG^{AP}						
dCTP	167 ± 15	1.03 ± 0.03	$6.2x10^{-3}$	9.2	-	98.4
dATP	856 ± 184	$(7.6 ± 0.8)x10^{-2}$	$8.9x10^{-5}$	0.2	$1.4x10^{-2}$	1.4
dGTP	955 ± 160	$(2.0 ± 0.2) x 10^{-3}$	$2.1x10^{-6}$	21	$3.3x10^{-4}$	0.0
dTTP	557 ± 36	$(5.5 ± 0.1)x10^{-3}$	$9.9x10^{-6}$	8.4	$1.6x10^{-3}$	0.2
*21/26-mer-dG^{AP}						
dGTP	674 ± 231	$(2.8 ± 0.3)x10^{-2}$	$4.2x10^{-5}$	88	-	88.6
dATP	886 ± 145	$(2.7 ± 0.2)x10^{-3}$	$3.1x10^{-6}$	0.8	$6.8x10^{-2}$	6.5
dCTP	328 ± 79	$(6.8 ± 0.5)x10^{-4}$	$2.1x10^{-6}$	15	$4.7x10^{-2}$	4.4
dTTP	2300 ± 461	$(4.9 ± 0.7)x10^{-4}$	$2.2x10^{-7}$	8.2	$5.1x10^{-3}$	0.5
22/26-mer-dG^{AP}						
dCTP	682 ± 80	6.3 ± 0.3	$9.3x10^{-3}$	3.7	-	99.8
dATP	826 ± 85	$(4.2 ± 0.2)x10^{-3}$	$5.0x10^{-6}$	0.8	$5.4x10^{-4}$	0.1
dGTP	502 ± 98	$(4.3 ± 0.3)x10^{-3}$	$8.6x10^{-6}$	0.8	$9.3x10^{-4}$	0.1
dTTP	1540 ± 257	$(4.7 ± 0.5)x10^{-3}$	$3.1x10^{-6}$	3.6	$3.3x10^{-4}$	0.0
23/26-mer-dG^{AP}						
dGTP	62 ± 18	1.9 ± 0.1	$3.1x10^{-2}$	0.8	-	98.9
dATP	668 ± 173	$(3.7 ± 0.4)x10^{-2}$	$5.5x10^{-5}$	0.3	$1.8x10^{-3}$	0.2
dCTP	1130 ± 161	$(9.0 ± 0.7)x10^{-3}$	$7.9x10^{-6}$	0.2	$2.6x10^{-4}$	0.0
dTTP	1290 ± 218	$(3.4 ± 0.3)x10^{-2}$	$2.7x10^{-4}$	0.01	$8.7x10^{-4}$	0.9

[a]Calculated as $(k_p/K_{d,\ dNTP})_{control}/(k_p/K_{d,\ dNTP})_{damaged}$.
[b]The values for $(k_p/K_{d,\ dNTP})_{control}$ are from Supplementary Table 1.
[c]Calculated as $(k_p/K_{d,\ incorrect\ dNTP})_{damaged}/[(k_p/K_{d,\ correct\ dNTP})_{damaged} + (k_p/K_{d,\ incorrect\ dNTP})_{damaged}]$
[d]Calculated as $\{(k_p/K_{d,\ dNTP})_{damaged}/[\Sigma(k_p/K_{d,\ dNTP})_{damaged}]\}x100$.
*Denotes pause sites.

By combining the measured k_p and $K_{d,\ DNA}$, the dNTP incorporation efficiency ($k_p/K_{d,DNA}$), efficiency ratio (relative to undamaged DNA), fidelity, and probability (Table 1.3) were calculated. Comparing the kinetic data for non-pause sites, the $k_p/K_{d,DNA}$ values for the correct dNTP incorporation were within 1-4-fold of the values for the control 26-mer and were 100-

4000-fold greater relative to misincorporations (Table 1.3). However, for the pause sites, 20/26-mer-dG^{AP} and 21/26-mer-dG^{AP}, the correct dNTP incorporation efficiencies, respectively, decreased by 9- and 88-fold (Table 1.3) in comparison to the corresponding control DNA substrates. Thus, the presence of the dG^{AP} lesion unfavorably impacted the correct incorporation both opposite the lesion and in the extension of the bypass product.

The kinetic parameters also provided insight into the fidelity of DNA synthesis at each position along the damaged DNA template. The fidelity ranged from 10^{-3} to 10^{-5}, upstream and downstream of the pause sites, which is similar to Dpo4'sfidelity replicating control DNA. In contrast, the fidelity at the pause sites, especially the 2nd pause site, was lowered by 10-100-fold compared with non-pause sites and with the control DNA. Based on the dNTP incorporation efficiency values in Table 1.3, Dpo4 catalyzed the insertion of dNTPs with the following preference: dCTP>>dATP>dTTP, dGTP at the first pause site and dGTP>dATP,dCTP>dTTP at the second pause site.

Biphasic Kinetics of dNTP Incorporation at the Pause Site- A colleague in the lab investigated if incorporation of dNTPs at the pause sites demonstrated biphasic kinetics, which would be hidden in the previously performed single-turnover dNTP assays. In order to separate the multiple kinetic phases, a DNA trap was added to the assay. For these assays, a 5 μM concentration of undamaged D-1 (Table 1.1) was used as a trap. To identify biphasic kinetics, a pre-incubated solution of Dpo4 (120 nM) and 5'-^{32}P-labeled DNA (30 nM) was rapidly mixed with a solution containing the correct dNTP (1.2 mM) and unlabeled D-1 DNA substrate(5 μM) for various times before terminating the reaction by quenching with 0.37 M EDTA. The time courses (Figure 1.6) of correct dNTP incorporation into 20/26-mer-dG^{AP} and 21/26-mer-dG^{AP},

the positions of the two pause sites, were both biphasic and were fit to Equation 1.4, which yielded the biphasic kinetic parameters listed in Table 1.4.

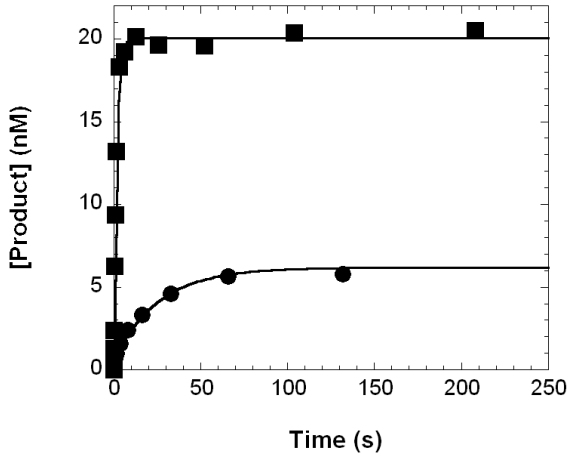


Figure 1.6 Biphasic kinetics of correct nucleotide incorporation in the presence of a DNA trap. The product concentration was plotted as a function of reaction time for each DNA substrate which was then fit to Eq 1.4. 20/26-mer-dG^{AP} (▪) or 21/26-mer-dG^{AP} (●).

Table 1.4. Biphasic kinetic parameters for correct dNTP incorporation into 5′-[^{32}P]-labeled DNA (30 nM) in the presence of a DNA trap (5 μM) at 37 °C.

DNA Substrate	Correct dNTP	A_1 (nM)	k_1 (s^{-1})	A_2 (nM)	k_2 (s^{-1})
19/26-mer	dGTP	27 ± 1 (91%)*	3.2 ± 0.4		
20/26-mer	dCTP	26 ± 1 (86%)*	4.6 ± 0.5		
21/26-mer	dGTP	27.0 ± 0.4 (88%)*	2.3 ± 0.1		
19/26-mer-dG^{AP}	dGTP	27.0 ± 0.4 (89%)*	2.5 ± 0.2		
20/26-mer-dG^{AP}	dCTP	2.0 ± 0.7 (6.7%)*	11 ± 6	18 ± 1 (60%)*	0.7 ± 0.1
21/26-mer-dG^{AP}	dGTP	0.9 ± 0.1 (3%)*	1.9 ± 0.3	5.7 ± 0.2 (19%)*	0.031 ± 0.003

*Calculated as (reaction amplitude/30 nM) x 100.

For the two pause sites, the rate of the first phase (k_1) was faster than the rate of the second phase (k_2), though the reaction amplitude of the first phase (A_1) was smaller than the amplitude of the second phase (A_2). Moreover, the total amplitudes ($A_1 + A_2$) at these sites were much less than 100% product formation. For the 20/26-mer-dG^{AP}, the total amplitude was 66.7%, while the 21/26-mer-dG^{AP} had a total amplitude of 22%. The control substrates and the 19/26-mer-dG^{AP} demonstrated only monophasic kinetics and had reaction amplitudes close to 90%. Thus, it was demonstrated through these DNA trap assays that the dG^{AP} lesion only altered the kinetics of dNTP incorporation at the two critical steps of translesion synthesis.

Discussion

The full length product seen in the running start assay (Figure 1.1B) demonstrated that Dpo4 is capable of bypassing a site-specifically placed dG^{AP} lesion. However, the initial formation of this full-length product was significantly slower with the damaged template, 26-mer-dG^{AP}, (180 s) than with the control template, 26-mer, (10 s). Even yet, the accumulation of intermediates 20- and 21-mer indicated pausing of Dpo4 when incorporating dNTPs opposite the lesion and in the extension of the bypass product. To determine the kinetic impact of the dG^{AP} lesion on DNA binding and dNTP incorporation at positions upstream, opposite, and downstream of the lesion, we utilized EMSA and pre-steady-state kinetic methods.

A Kinetic Basis for the Pausing of Dpo4 in the Presence of dG^{AP}- Figure 1.1B demonstrated the accumulation of two intermediates, the 20- and 21-mer, which indicated a

stalling of Dpo4 at these sites on the damaged template. This incorporation profile can be explained through a comparison of the catalytic efficiencies of the positions leading up to full product formation, where the accumulation of 25-mer is considered full length product. For example, incorporation of the correct dGTP into 21/16-mer-dG^{AP} is 150-fold less efficient than correct dCTP into 20/26-mer-dG^{AP}, which is subsequently 4-fold less efficient than correct dGTP into 19/26-mer-dG^{AP}. These inefficiencies of incorporation led to the accretion of 20- and 21-mer with the later accumulating more than the former. In contrast, incorporation of dGTP into the 23/26-mer-dG^{AP} was 3-fold less efficient than incorporation of dCTP into the 22/26-mer-dG^{AP}, which was 220-fold more efficient than incorporation of dGTP into the 21/26-mer-dG^{AP}. Thus, no accumulation was found for the 22- and 23-mer. In general, if a polymerase stalls at a site, its elongation is less efficient than its production. The contrary is true of intermediates that do not show pausing through the accumulation at a nonproduct position.

The efficiency ratio compares the incorporation efficiencies of the control over the damaged substrates. As previously stated these substrate specificities combine the values of the k_p and the $K_{d,dNTP}$. Thus, to determine which microscopic kinetic parameter contributed to the accumulation of the intermediate DNA substrate, the k_p and the $K_{d,dNTP}$ for each site along the DNA template were compared to values of the control at the corresponding position (Figure 1.7 and Figure 1.8).

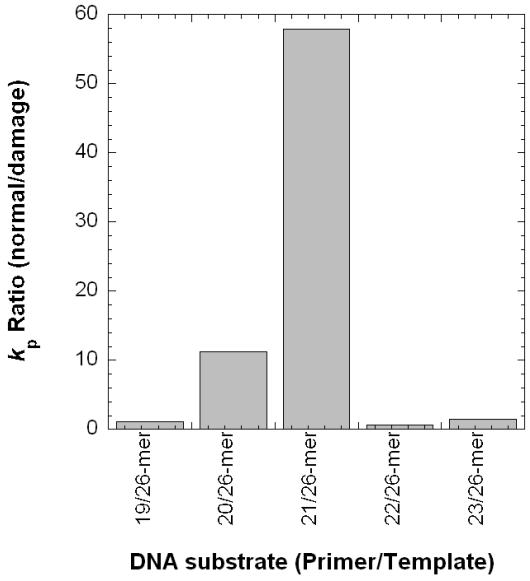


Figure 1.7: dG^{AP} lesion's effect on k_p of correct nucleotide incorporation catalyzed by Dpo4.

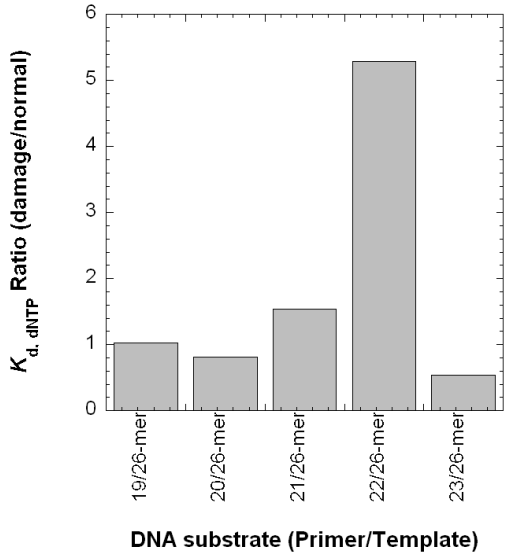


Figure 1.8: dG^{AP} lesion's effect on $K_{D,\ DNA}$ of correct nucleotide incorporation catalyzed by Dpo4.

Figure 1.7 demonstrates that the k_p was significantly affected only at the two pause sites (11-fold for 20/26-mer-dG^{AP} and 58-fold for 21/26-mer-dG^{AP}) though, the $K_{d,dNTP}$ values were within 5-

fold of the control (Figure 1.8). Thus, the inefficiency of dNTP incorporation at the pause sites was primarily due to a decrease in the k_p.

DNA trap assays were performed to determine if the slow k_p values were a result of DNA being ensnared in nonproductive complexes with Dpo4. The results of the trap assay supported this hypothesis in that the pause sites demonstrated biphasic kinetics, whereby a small, fast phase (A_1 and k_1) proceeded a large, slow phase (A_2 and k_2). The combination of the two phases distinguished in the DNA trap assay were close to the calculated k_{obs} estimated in the single turnover experiments for both the 20- and 21/26-mer-dG^{AP}. In contrast, the DNA trap experiments with 19/26-mer-dG^{AP} revealed only one fast kinetic phase for the incorporation of dGTP. This data suggests that the dG^{AP} lesion did not affect nucleotide incorporation at a non-pause site and that both 19/26-mer-dG^{AP} and the control substrates were bound to Dpo4 in a productive manner. In comparison, the small fast phases of dNTP incorporation with both the 20/26-mer-dG^{AP} and the 21/26-mer-dG^{AP} occurred with similar rates (k_1, Table 1.4) to the single-turnover rates observed with the control DNA substrates (Table 1.3). Thus, small percentages (A_1) of these DNA substrates were bound productively by Dpo4 (E·DNA$^{P}_{n}$) in the fast phase. In the slow phase, larger percentages (A_2) of the DNA substrates must have been bound in a less catalytically efficient mode by Dpo4 (E·DNA$^{N}_{n}$), because they were elongated with much slower rates (k_2). Moreover, the elongation of E·DNA$^{N}_{n}$ occurred in a single binding event, demonstrating that a slow conversion from E·DNA$^{N}_{n}$ to E·DNA$^{P}_{n}$ must take place, followed by a rapid extension (k_1). The total reaction amplitudes of the 20/26-mer-dG^{AP} (66.7%) and the 21/26-mer-dG^{AP} (22%) had much smaller reaction amplitudes than the amplitudes of the 19/26-mer-dG^{AP} and control substrates (≈90%). The partitioning between the dissociation of E·DNA$^{N}_{n}$ and the conversion of E·DNA$^{N}_{n}$ to E·DNA$^{P}_{n}$ led to a reduction in A_2 by a factor of $k_2/(k_2 + k_{off})$,

where k_{off} is the DNA dissociation rate from Dpo4·DNA. The k_{off} of D-1 (Table 1.1) was previously determined to be 0.02 s^{-1}. This value was used for the k_{off} of both the 20/26-mer-dG^{AP} and the 21/26-mer-dG^{AP} because Dpo4 bound to these substrates with tighter or similar affinities as control DNA (Table 1.2). Using the estimated and determined k_{off}, k_2, A_1, and A_2 values, we estimated that 62% of 20/26-mer-dG^{AP} and 31% of 21/26-mer-dG^{AP} are in the form of $E{\cdot}DNA^N_n$ and that 21.3% of 20/26-mer-dG^{AP} and 56% of 21/26-mer-dG^{AP} were never elongated. Thus, significant amounts of 20/26-mer-dG^{AP} and 21/26-mer-dG^{AP} were bound by Dpo4 in a catalytically incompetent mode ($E{\cdot}DNA^D_n$). Overall, these DNA trap experiments suggest a kinetic mechanism for bypassing dG^{AP} as shown in Figure 1.9.

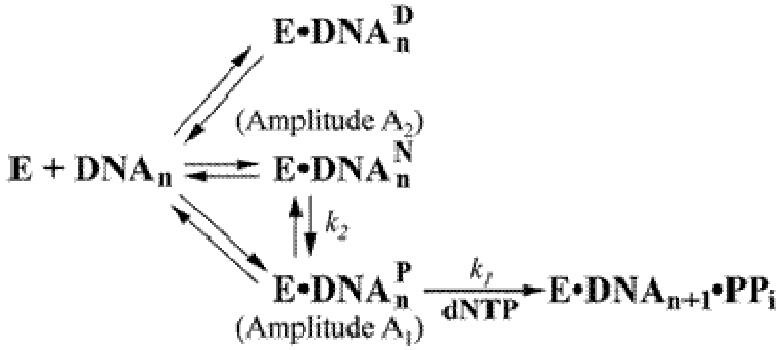


Figure 1.9: Proposed kinetic mechanism for the bypass of dG^{AP} catalyzed by Dpo4.

This proposed mechanism for lesion bypass is supported by existing structural evidence. For example, the combined NMR molecular mechanics computational studies reveal that the dG^{AP} moiety of an embedded dG^{AP}:dC base pair in an 11-mer duplex is only intercalated into the DNA helix between adjacent Watson-Crick base pairs [24]. Due to the syn glycosidic torsion angle, both bases of the dG^{AP}:dC base pair are displaced into the major groove [24]. If dG^{AP} at

the pause sites maintains this conformation, then the dG^{AP} lesion will occupy the position of the incoming dNTP, thereby blocking catalysis. This binary complex could not undergo synthesis without undergoing dramatic structural changes and likely represents the $E{\cdot}DNA^{D}_{n}$. Other studies using x-ray crystallography, suggest the presence of other conformations in which the 1-AP is quasi-intercalative or externally bound [11, 25].In one study, the AAF of damaged DNA is situated within the major groove open pocket and pairs with the dCTP with modest polymerase perturbation [26]. This conformation seems to indicate that the translocation of the little finger of Dpo4 would be hindered. In another study of BPDE, there are two conformations of BPDE, one intercalated between base pairs and another flipped out of the DNA helix into a gap between the little finger and core domains [11, 25]. When the BPDE is intercalated, the distance between the 3'-OH of the primer is 9 Å, whereas when the BPDE is flipped out the distance is 3.9 Å. The optimum distance for catalysis is 3.4 Å [25]. Thus, extending this analysis to the dG^{AP} studies, if the dG^{AP} is flipped out the molecules of Dpo4·20/26-mer-dG^{AP} will be in the form of $E{\cdot}DNA^{P}_{n}$. Contrastingly, if the dG^{AP} is quasi-intercalated, the same binary complex will require structural change to establish a conformation in which efficient catalysis can occur, and thus is likely in the form $E{\cdot}DNA^{N}_{n}$. However, to verify and clarify this structural analysis, a collaborator of the Suo laboratory is working to crystallize the structure of Dpo4·DNA-dG^{AP}.

Unique Enhancement of DNA Binding- Interestingly, our EMSA studies of DNA binding reported in Table 1.2 showed that Dpo4·20/26-mer-dG^{AP} is about 4-fold tighter than Dpo4·21/26-mer-dG^{AP}, and Dpo4·20/26-mer is the only binary complex affected by the lesion. This increased binding suggests that the dG^{AP} in the Dpo4·20/26-mer-dG^{AP} likely interacts with the residues of the little finger domain of Dpo4, which correlates with the structures of Dpo4 and BPDE-dG [11]. Usually DNA lesions tend to distort the DNA structure and weaken the binding

of DNA to a DNA polymerase [10]. However, the presence of the dG^{AP} lesion tightened the binding between Dpo4 and the DNA substrate, which may explain why the 20/26-mer-dG^{AP} was elongated with 150-fold higher efficiency than 21/26-mer-dG^{AP}.

The Effect of dG^{AP} on dNTP incorporation at Adjacent Sites- The kinetic data found in Table 1.3 shows that dG^{AP} did not kinetically affect dNTP incorporation at any downstream position of the pause sites. In contrast, both an abasic site and a cisplatin-d(GpG) adduct kinetically affected six to seven downstream dNTP incorporation during translesion synthesis catalyzed by Dpo4 [8, 10]. These differences could be a reflection of how each lesion distorts the structure of DNA within a DNA polymerase active site.

Final Conclusions about the Kinetic Mechanism of DNA Lesion Bypass- The bisphasic kinetics of dNTP incorporation at pause sites have been observed in the bypass of an abasic site catalyzed by Dpo4 [8]as well as a cisplatin-d(GpG) adduct catalyzed by Dpo4 [10]. As with the bypass of dG^{AP}, the total reaction amplitude in each of these cases is much less than the reaction amplitude obtained with either a control DNA or a DNA substrate at a non-pause site, indicating the existence of the $E{\cdot}DNA^{D}_{n}$ form. Some of these dead end binary complexes can bind a nucleotide and form dead-end binary complexes.

In conclusion, we have determined that Dpo4 is capable of bypassing a model single-base lesion dG^{AP} in a kinetically inefficient and error prone manner. The extension step rather than insertion opposite the dG^{AP} lesion was more challenging for the enzyme. Through this study, we have established the first kinetic mechanism for the dG^{AP} lesion bypass.

Chapter 2: Determination of the Mutagenic Profile of Y-family DNA Polymerases Bypassing a 1-Nitropyrene Induced Lesion

Background and Introduction

The recently discovered Y-family DNA polymerases have been shown to function primarily in the bypass of replication stalling DNA lesions. In overcoming DNA lesions, these DNA polymerases reduce the possibility of cell apoptosis induced by DNA damage. Moreover, the significance of these Y-family DNA polymerases is underscored by their presence in organisms in all three domains of life. One such Y-family DNA polymerase is *Sulfolobus sulfataricus's* DNA polymerase IV (Dpo4), which was described in depth in Chapter 1. While *Sulfolobus sulfataricus* has only one Y-family enzyme, in humans, four Y-family DNA polymerases have been identified. These polymerases are known as human DNA polymerases eta (hPolη), iota (hPolι), kappa (hPolκ), and Rev1. Like most Y-family DNA polymerases, these enzymes are capable of both error-free and error-prone translesion synthesis *in vivo*. However this process is dependent on the specific lesion present in the DNA. For example, hPolη is responsible for the error-free bypass of *cis-syn* thymine-thymine (TT) dimers [27, 28]. Genetic mutational inactivation of hPolη leads to a cancer-prone syndrome known as xeroderma pigmentosum variant (XPV) that predisposes individuals to an increased incidence of sunlight-induced skin cancer [28]. Besides the *cis-syn* TT dimers, hPolη has been shown to bypass a variety of DNA lesions, including abasic sites (AP sites) [29], 8-ozoguanini (8-oxoG) [29], (+)-trans-anti-benzo[a]pyrene-N^2-dG ((+)BPDE-dG) [29], 1,N^6-Ethenodeoxyadenosine [30], O^6-methylguanine [31], N-2-acetylaminofluorene-dG (AAF-dG) [32], and cisplatin-dGpG

intrastrand adducts [32]. In contrast, hPolι is the most error-prone human Y-family DNA polymerase known to date, but has been shown to traverse AP sites [33], 8-oxoG,AAF-dG[34] , *cis-syn* TT dimmers [35], and other TT photoproducts [36]. hPolκ is a member of the DinB subfamily and is the closest of the human Y-family DNA polymerases to Dpo4 [37]. hPolκ has been shown to bypass AP sites , 8-oxoG, AAF-dG, and (+)BPDE-dG [38]. Finally, Rev1, which is the only Y-family DNA polymerase to contain a BRCT domain, has been implicated in the mediation of protein-protein interactions and is classified as a dCTP transferase [39]. Human Rev1 has been shown to incorporate dCTP opposite AP sites [40], 8-oxoG [41], and (+)BPDE-dG [41]. Based on the published studies, there is clearly significant overlap of the *in vitro* lesion bypass spectra for the four human DNA polymerases. However, it is still unclear which Y-family DNA polymerase bypasses specific lesions *in vivo*.

In chapter one, a specific lesion known as *N*-(deoxyguanosin-8-yl)-1-aminopyrene ($dG^{AP)}$ which forms from a reaction between DNA and the metabolites of 1-nitropyrene (1-NP), was described and investigated. The model Y-family DNA polymerase Dpo4 was shown capable of bypassing this lesion and the mechanism and fidelity of this bypass were determined. However, to extend the study of the effect of the dG^{AP} lesion on replication by Y-family DNA polymerases, we set out to gather DNA sequence data for Dpo4, as well as for two of the human Y-family DNA polymerases. While kinetic analysis can provide insight into the fidelity of these enzymes, it cannot capture the full mutagenic profile of the Y-family DNA polymerase's bypass of the dG^{AP} lesion. For example, kinetic analysis cannot capture infrequent mutation events or deletions and additions that could occur during dG^{AP} bypass. Thus, we exploited a novel experiment developed in our lab, short oligonucleotide sequencing assay (SOSA) [42], to

determine the mutational spectra of three Y-family DNA polymerases, Dpo4, hPolκ, and hPolη .

Methods

Materials- T4 DNA Ligase and Taq DNA polymerase were purchased from Fermentas and Invitrogen, respectively. [γ-^{32}P]ATP was purchased from GE Healthcare Life Sciences. The oligonucleotides listed in Table 2.1 were purchased from Integrated DNA Technologies, and were purified, labeled, and annealed as previously described [15]. The TOPO TA Cloning Kit for Sequencing was purchased from Invitrogen. Its contents were used in both the PCR and transformation steps.

Table 2.1. Sequences of DNA oligonucleotides

Primers	
17-merS	5′-CTACCTGAACGACGGCC-3′
18-merS	5′-GGTCAACATTGCACTAGC -3′
Templates	
26-mer	3′-TTGCTGCCGGTCACTTAAGCGCGCCC-5′
[a]26-mer-dG^{AP}	3′-TTGCTGCCGGTCACTTAAGC**G**CGCCC-5′
73-mer	3′-CTACTCAGCCGTTGATGGACTTGCTGCCGGTCACTTA AGCGCGCCCCTGTCCTGCCGATCACGTTACAACTGG-5′
[a]73-mer-dG^{AP}	3′-CTACTCAGCCGTTGATGGACTTGCTGCCGGTCACTTA AGC**G**CGCCCCTGTCCTGCCGATCACGTTACAACTGG-5′

[a]**G** designates *N*-(deoxyguanosin-8-yl)-1-aminopyrene (dG^{AP}).

Protein Purification- My colleagues in Dr. Suo's laboratory purified the proteins used for this study, which included *Sulfolobus sulfataricus* Dpo4, hPolη, and hPolκ. The full length Dpo4 was expressed in *E. coli* and purified within as previously described [7].

For DNA polymerase η, the gene encoding the C-terminal His-tagged enzyme was cloned

into the *Nde*I and *Xho*I sites of pET21B to generate a plasmid pET-21B-hPolη. The C-terminal His_6-tagged hPolη was induced and expressed in *E. coli* BL21(DE3) Rosseta cells at 16 °C. The protein was purified through a nickel affinity column, a heparin sepharose column, and a HiTrap SP column. The concentration of purified hPolη was measured spectrophotometrically at 280 nm using the calculated extinction coefficient of 70731 $M^{-1}cm^{-1}$.

The gene encoding a truncated fragment (residues 9-518) of hPolκ (hΔPolκ) was subcloned into the *Nco*I and *Xho*I sites of a plasmid pHIS-Parallel1 (24) to create pHIS-hPolkappa-9-518. This plasmid was transformed into *E. coli* BL21(DE3) Rosetta cells. The N-terminal His_6-tagged hΔPolκ was induced and expressed at 19 °C, and then was purified through a nickel affinity column, a heparin sepharose column, a HiTrap Q column, and a Sephacryl 200 gel filtration column. The concentration of purified hΔPolκ was measured spectrophotometrically at 280 nm using the calculated extinction coefficient of 31860 $M^{-1}cm^{-1}$.

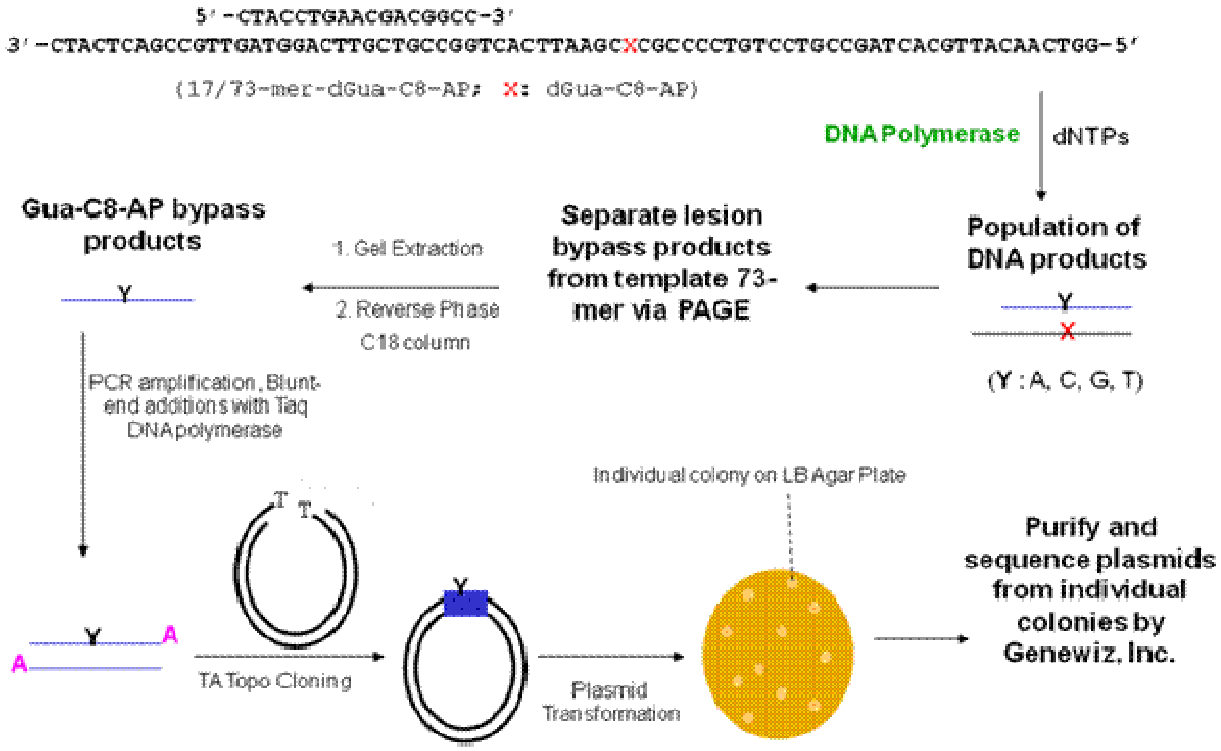
Buffer- The reaction buffer R contained 50 mM HEPES (pH 7.5 at 37 °C), 5 mM $MgCl_2$, 50 mM NaCl, 5 mM DTT, 10% glycerol, 0.1 mM EDTA, and 0.1 mg/ml BSA. All reactions were performed at 37 °C unless otherwise indicated.

Running Start Assays- A running start assay was performed as previously described for each Y-family DNA polymerase studied [7, 10] . In brief, a solution of 5'-^{32}P-labeled DNA(17/26-mer-dG^{AP} or 17/26-mer)(100 nM) and DNA polymerase (100 nM) in buffer R was pre-incubated at 37 ˚C and subsequently rapidly mixed with a solution containing all four dNTPs (200 μM each) via a rapid chemical-quench flow apparatus (KinTek). The reaction was quenched with .37 M EDTA after various reaction times. The reaction products were visualized

by denaturing PAGE (17% polyacrylamide, 8 M urea).

Short Oligonucleotide Sequencing Assay (SOSA)- Each Y-family DNA Polymerase (120 nM) was pre-incubated with the DNA substrate (17/73-mer-dG^{AP} or 17/73-mer)(30 nM, Table 1) and subsequently mixed with all four dNTPs (200 μM each) for an optimized time period at 37 °C. The reaction time was optimized to produce sufficient full length bypass product (60-mer) or near full length bypass product. To terminate the reaction, the solution was treated with three cycles of freeze-thaw (-196 °C and 95 °C respectively) denaturation. SOSA was performed using a control DNA substrate (17/73-mer) and Dpo4 to examine the quality of the method and to compare the mutagenic profile of Dpo4 under normal replications and in translesion synthesis. The mutagenicity of Dpo4, hpol η, and hΔpol κ's bypass of the dG^{AP} lesion were determined by performing the SOSA with the 17/73-mer-dG^{AP}. This DNA template (73-mer-dG^{AP}) contained a dG^{AP} lesion 33 positions from the 5' end (Table 1). After generating a population of "dG^{AP} bypass product", these products were isolated using PAGE, in which radiolabeled markers were utilized to determine the location of full length or near full length product. The isolated products were purified as previously described [22]. Once purified, the products were amplified using polymerase chain reaction (PCR), utilizing *Taq* DNA polymerase, 17-merS and 18-merS (Table 2.1). These primers were designed to provide sequence information for 10 nucleotides upstream from the dG^{AP} lesion and 12 nucleotides downstream from the lesion in addition to incorporation events directly opposite of the dG^{AP} lesion. Single dA's were placed on the blunt ends of the double stranded DNA by *Taq* DNA polymerase during the PCR step. After amplification, the population of "bypass product" was ligated into a pCR4-TOPO vector using a TOPO TA cloning kit (Invitrogen) and subsequently transformed into a TOP10 strain of *E. coli* (Invitrogen) as previously described [42]. The DNA plasmid containing the ligated bypass product was

sequenced from bacterial colonies (Genewiz, South Plain, NJ). This method is summarized in Scheme 2.1.



Scheme 2.1: SOSA Overview

Results

Running Start Assays of three Y-family DNA polymerases-While full-length Dpo4 and hPolη were utilized, recombinant hΔPolκ (residues 9-518) was used as the enzyme, rather than its full-length protein. For the running start assays, these enzymes were expressed and purified from *E. coli* as soluble and active proteins.

The running start assays were performed for each Y-family DNA polymerase under the same reaction conditions as described in the Materials and Methods. The results for Dpo4's

elongation past the dG^{AP} adduct were seen in Figure 1.1 and described in the Chapter 1 Results section. In brief, Dpo4 synthesized full-length product 26-mer with the undamaged DNA substrate 17/26-mer (Table 2.1) within 10 s (Figure 1.1A). However, while Dpo4 was capable of synthesis bypassing the dG^{AP} lesion, the enzyme was significantly slowed and full-length product was not observed until after 180 s (Figure 1.1B). Interestingly, during synthesis, an accumulation of intermediate products, 20- and 21-mer, indicated that Dpo4 stalled in its dNTP incorporation directly opposite the dG^{AP} lesion and in the extension of the lesion bypass product, respectively.

The running start assay for hΔPolκ generated similar results to those of Dpo4. hΔPolκ synthesized near full-length product (25-mer) with the undamaged DNA substrate 17/26-mer (Table 2.1) within 3 s (Figure 2.1A). However, in bypassing the dG^{AP} lesion, hΔPolκ was significantly slowed and the 25-mer was not observed until 30 s (Figure 2.1B). Accumulation of the 25-mer rather than the 26-mer in Figure 2.1A and Figure 2.1B was likely caused by polymerase "slippage" via primer realignment. Moreover, like Dpo4, the accumulation of intermediate products, 20- and 21-mer, indicated stalling in dNTP incorporation directly opposite the dG^{AP} lesion and in the extension of the lesion bypass product, respectively. Again, no pausing was demonstrated in the DNA synthesis opposite the control template.

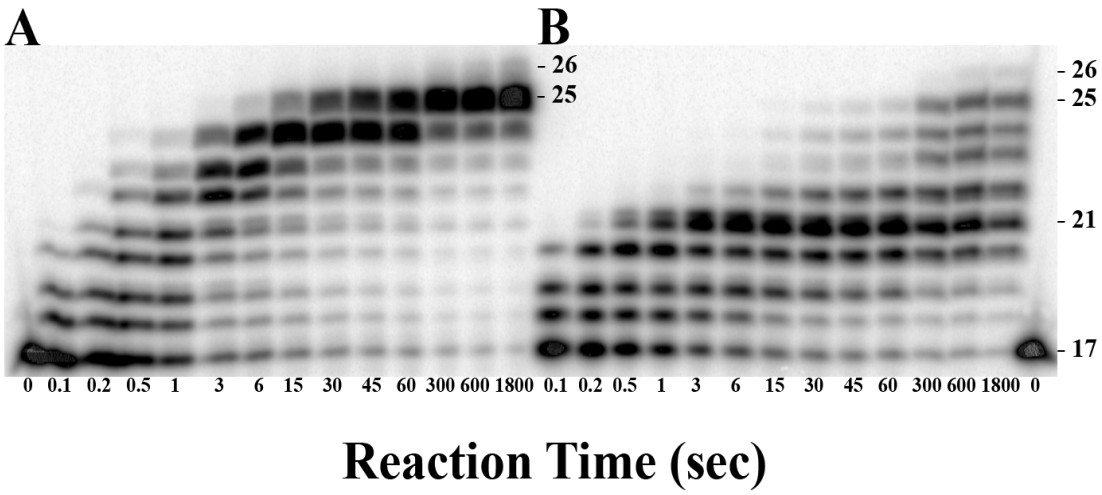


Figure 2.1: Running start assay for hΔPolκ. (A) 17/26-mer; (B) 17/26-mer-dG^{AP}. Sizes of important products are indicated, and the 21st position marks the location of the dG^{AP} lesion from the 3'-terminus of the DNA template.

In contrast to Dpo4 and hΔPolκ, the running start assay for hPolη showed the generation of full-length product (26-mer) in 3 s for both the of 5'-^{32}P-labeled control (17/26-mer) and damaged DNA substrate (17/26-mer-dG^{AP}) (Figure 2.2A and B). In Figure 2.2A and B, formation of the 27-mer was likely due to blunt end addition [23]. Even more interesting to the rate of bypass, hPolη does not seem to have the same pausing pattern as Dpo4 and hΔPolκ. In fact, hPolη does not seem to pause at all when bypassing and extending synthesis past the dG^{AP} lesion. However, in general, there do seem to be more of an accumulation of intermediates for the replication of the damaged template versus the control.

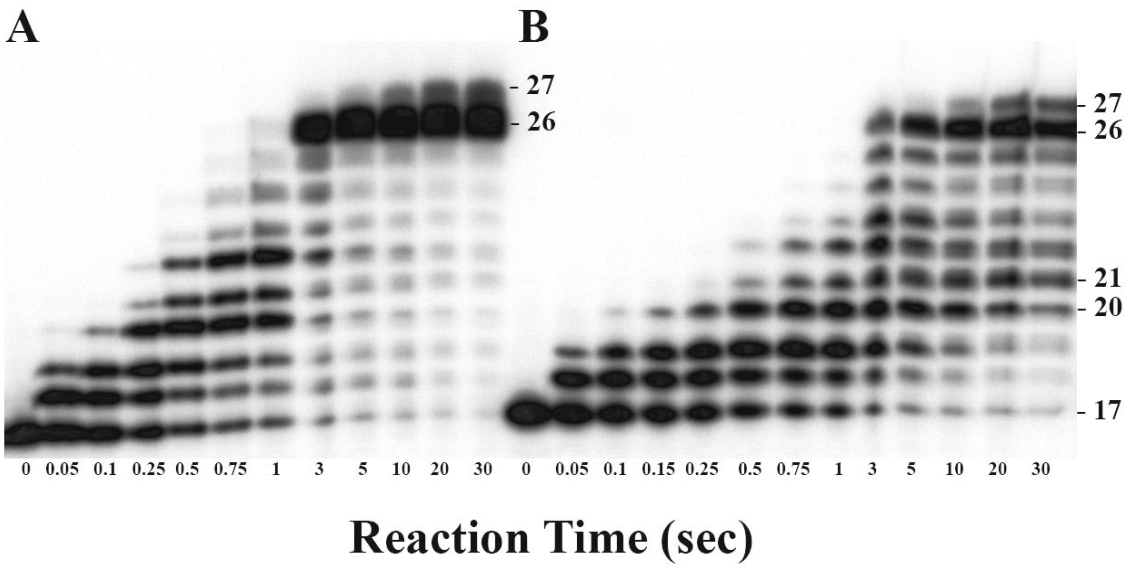


Figure 2.2: Running start assay for hPolη. (A) 17/26-mer; (B) 17/26-mer-dG^{AP}. Sizes of important products are indicated, and the 21st position marks the location of the dG^{AP} lesion from the 3'-terminus of the DNA template.

Thus, both Dpo4 and hΔPolκ were able to bypass the dG^{AP} . However, their rate of incorporation slowed when encountering this lesion and extending the bypass product. At these two pause sites, the incorporation profile indicated accumulation of intermediate products 20-mer and 21-mer, suggesting slow turnover at these sites in comparison to the corresponding assay performed with 17/26-mer. Furthermore, the accumulation at the second pause site was greater than that at the first pause site, indicating that it was more of an enzymatic challenge to extend past the incorporation opposite the lesion than to perform that incorporation itself. In contrast to Dpo4 and hΔPolκ, hPolη bypassed the dG^{AP} site relatively efficiently and subsequent downstream incorporation was not significantly perturbed by the lesion. However, while there was no obvious difference in the amount of time required for hPolη to generate the full-length product with the control and damaged template, in general, there was more accumulation of

intermediates for the damaged template.

These running start assays will be used to analyze the mutagenic profiles. However, the template used for the SOSA assay was the 17/73-mer-dG^{AP}. Thus, in the future the running start assays will be performed with the 17/73-mer-dG^{AP} in order to better compare the results with those of the SOSA assay.

Quantitative analysis of the mutagenic profiles of Dpo4 synthesis of the control template- In order to confirm the quality of the SOSA method and to ensure mutations were not due to sequence bias, SOSA was performed on undamaged 73-mer with Dpo4. Among the 52 colonies sequenced, there were only two single-base deletions and two base substitutions (Figure 2.3). The total error frequency (4 mutations/ (25 nucleotides x52 incorporation events)) of Dpo4 was calculated to be $3x10^{-3}$, which corresponds to the fidelity determined by pre-steady state kinetic assays in Chapter 1 (10^{-3}-10^{-5}).

Cytosine Incorporations

```
CTACCTGAACGACGGCCAGTGAATTCGCGCGGGGACAGGACGGCTAGTGCAATGTTGACC      (48/52)
                C                                                  (1/52)
                    C                                              (1/52)
                  G                                                (1/52)
```

No Incorporation

```
CTACCTGAACGACGGCCAGTGAATTCG-GCGGGGACAGGACGGCTAGTGCAATGTTGACC      (1/52)
```

Figure 2.3: Mutagenic profile for Dpo4's replication of 17/73-mer. Results from the SOSA are shown separately based on specific nucleotide incorporated by Dpo4 in the control template. Sequences corresponding to primers used for amplification are shown in lower-font while sequenced nucleotides are upper font and underlined. Individual base substitutions (blue), deletions (red), and additions (green) are shown. Cyan highlighted font indicates position where the lesion will be in the 17/73-mer-dG^{AP}.

Quantitative analysis of the mutagenic profiles of dG^{AP} lesion bypass catalyzed by Dpo4, hΔPolκ, and hPolη- Due to the fact that single nucleotide kinetic analysis of the incorporation events in the vicinity of the dG^{AP} lesion may not accurately reflect all of the possible catalytic transactions occurring in this region of the DNA substrate, we applied SOSA to determine the precise sequence of DNA synthesized by each Y-family DNA polymerase. Although the DNA synthesis rates for each enzyme vary, we optimized the time for each enzyme to generate the full-length bypass products and thus encompassed the entire bypass product, including the initial approach of the enzyme to the lesion, its bypass, and subsequent extension. The SOSA was designed to provide sequence information for 10 nucleotides upstream from the dG^{AP} lesion and 12 nucleotides downstream from the lesion in addition to incorporation events directly opposite of the dG^{AP}.

The sequencing spectra for Dpo4's bypass of the dG^{AP} lesion can be found in Figure 2.4. Results for Dpo4 revealed that 90% of the bypass products (47/52 colonies) are characterized by an incorporation of dCTP opposite the lesion. In addition, Dpo4 incorporated an incorrect dATP (2%) and generated a deletion at the site of the lesion (8%). As the running start assay indicated, there was significant stalling of Dpo4 not only when incorporating nucleotides opposite the lesion, but also in extending the product one base pair downstream of the lesion. Thus, it is important to analyze the mutagenic profile of this downstream position. For Dpo4, 90% of the bypass products contained a dGTP at the position one base pair downstream from the lesion. Mutations at the downstream position included two substitutions and two deletions. Besides the mutations corresponding to the pause sites of Dpo4, random mutations were generated both upstream and downstream of the lesion (Figure 2.4). Thus of the 90% of those sequences with a correct incorporation opposite the lesion, 23% included mutations, either

deletions or substitutions, primarily upstream of the lesion or at the site directly downstream of the lesion. Overall, not including mutations opposite the lesion site, 27% of the bypass products contained mutations.

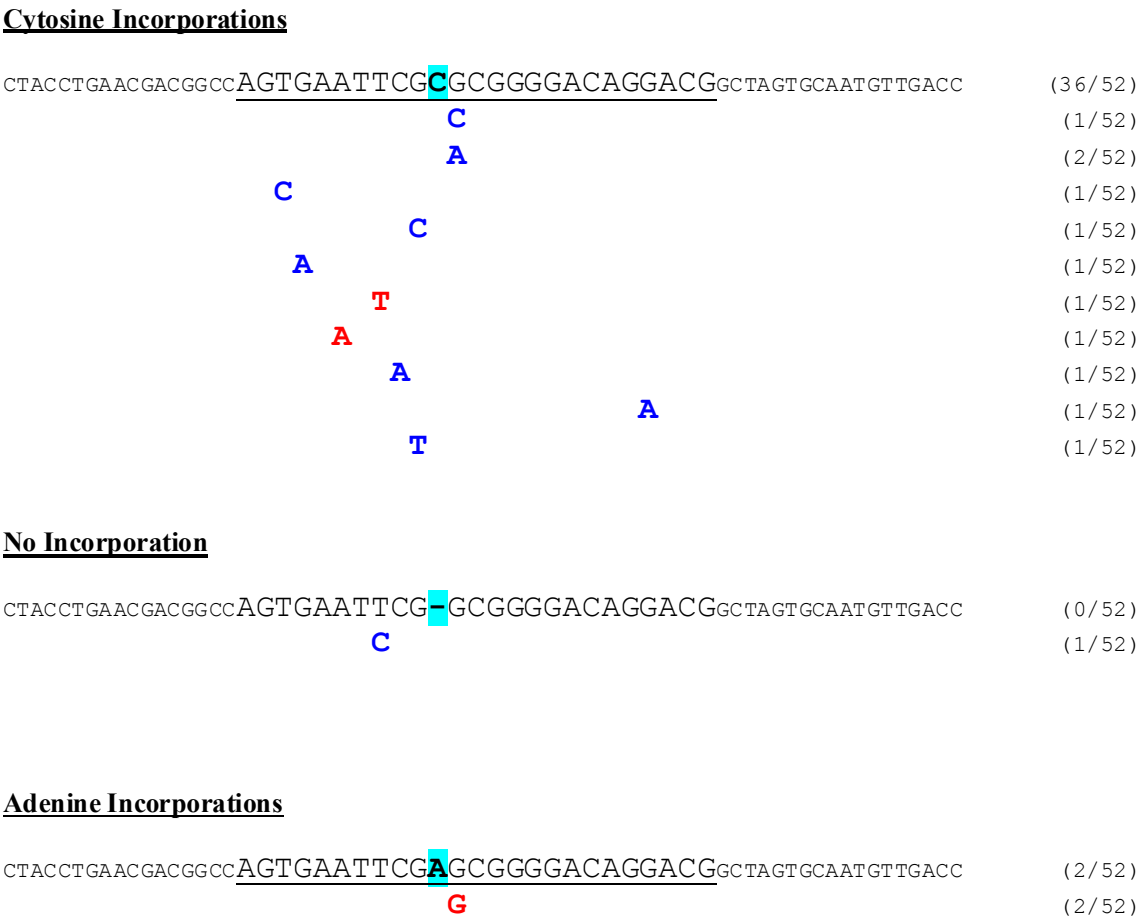
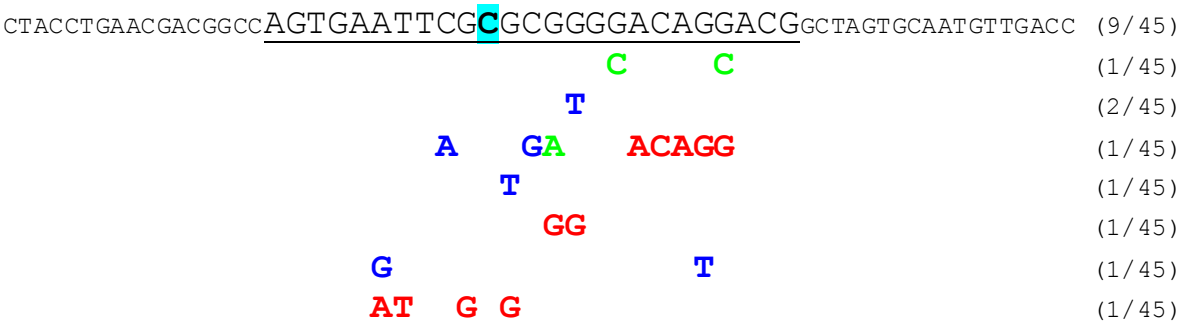


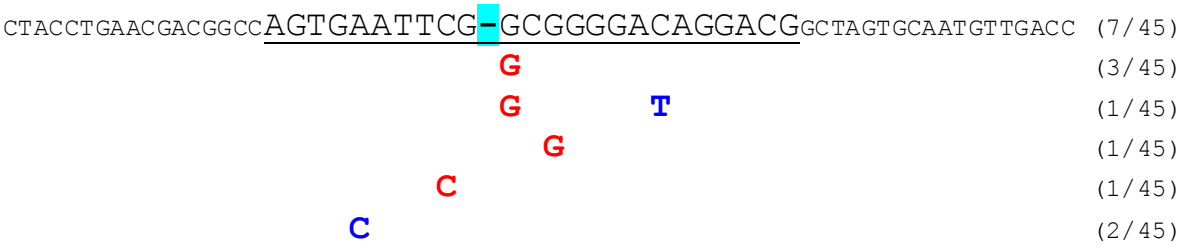
Figure 2.4: Mutagenic profile for Dpo4's replication of 17/73-mer-dG^{AP}. Results from the SOSA are shown separately based on specific nucleotide incorporated by Dpo4 in the control template. Sequences corresponding to primers used for amplification are shown in lower-font while sequenced nucleotides are upper font and underlined. Individual base substitutions (blue), deletions (red), and additions (green) are shown. Cyan highlighted font indicates position of the lesion.

As demonstrated in the sequencing spectra (Figure 2.5) hΔPolκ lacked a strong preference in incorporating nucleotides opposite the dG^{AP} lesion. The polymerase incorporated the correct dCTP (38%), an incorrect dATP (28%), an incorrect dTTP (2%), and caused a deletion at this site (33%). Moreover, hΔPolκ incorporated the correct dGTP in 84% in extending past the incorporation at the site of the lesion. Of the mutations produced at this site, 71% were deletions and 29% were substitutions. hΔPolκ's bypass products were also associated with mutations both upstream and downstream of this lesion (Figure 2.5). Not including stands that had a mutation only at the lesion site, 49% of the bypass strands from hΔPolκ had mutations.

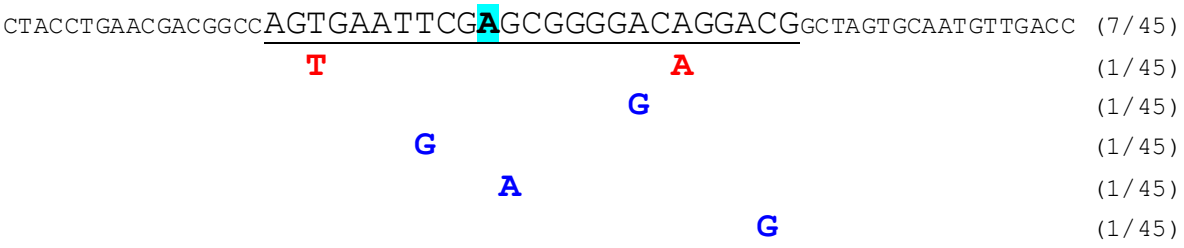
Cytosine Incorporations



No Incorporations



Adenine Incorporations



Thymine Incorporation

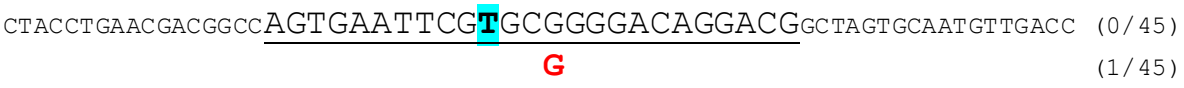
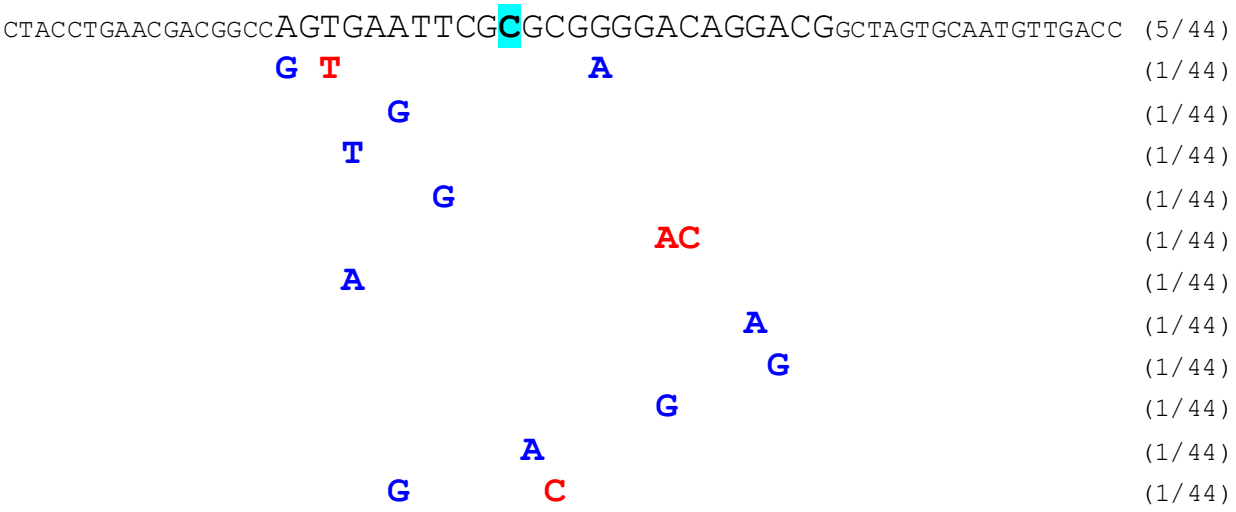


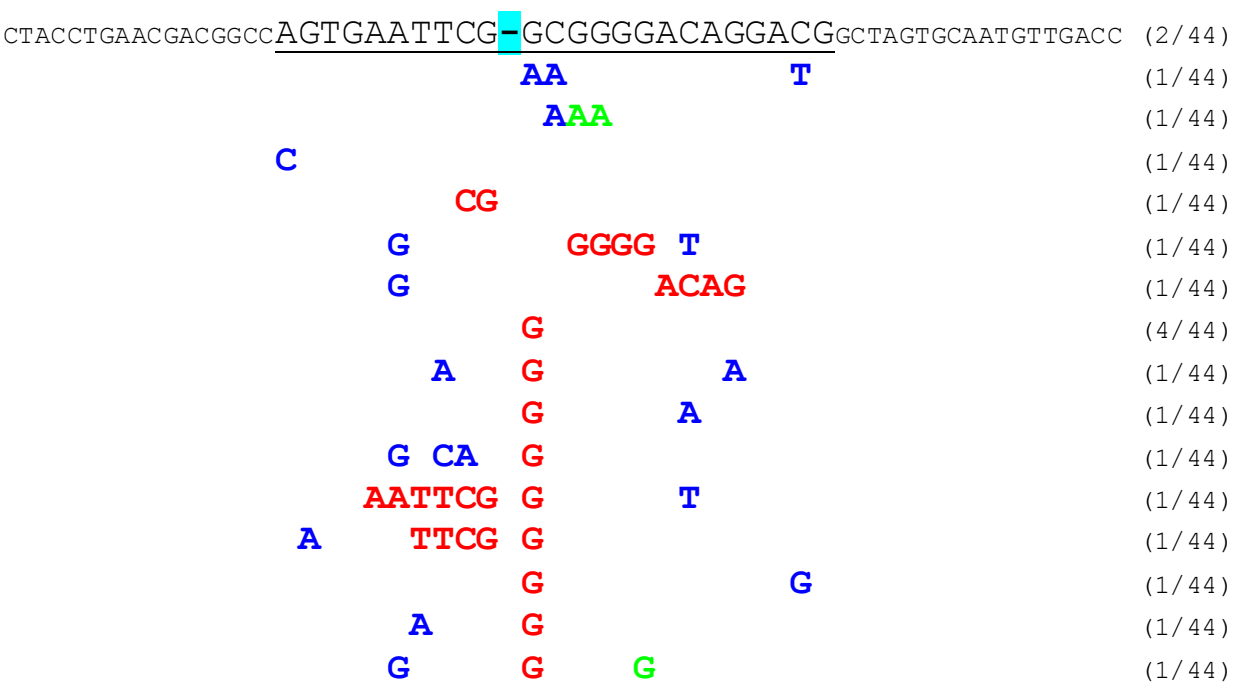
Figure 2.5: Mutagenic profile for hΔPolκ's replication of 17/73-mer-dG^{AP}. Results from the SOSA are shown separately based on specific nucleotide incorporated by Dpo4 in the control template. Sequences corresponding to primers used for amplification are shown in lower-font while sequenced nucleotides are upper font and underlined. Individual base substitutions (blue), deletions (red), and additions (green) are shown. Cyan highlighted font indicates position of the lesion.

Like hΔPolκ, hPolη lacked a strong preference in incorporating nucleotides opposite the dG^{AP} lesion. As represented in its sequencing spectra (Figure 2.6), the polymerase incorporated the correct dCTP (36%), an incorrect dATP (11%), an incorrect dTTP (5%), an incorrect dGTP (2%), and generated a deletion at the site opposite the lesion (45%). At the position directly downstream of the lesion, the correct dGTP was incorporated in 68% of the bypass products. Of those products that hPolη induced a mutation at this site, 86% were deletions and 14% were substitutions. hPolη also generated mutations both up and downstream of the lesion. Not including mutations at the site of the lesion, 80% of the hPolη bypass products had mutations. These mutagenized bypass products included some in which hPolη generated large deletions (4 or more consecutive nucleotides).

Cytosine Incorporations



No Incorporation



Adenine Incorporation

```
CTACCTGAACGACGGCCAGTGAATTCGAGCGGGGACAGGACGGCTAGTGCAATGTTGACC (1/44)
                                      C  GACG                 (1/44)
                    G                 G  GACG                 (1/44)
                           A          ACAG  C                 (1/44)
                              G                               (1/44)
```

Thymine Incorporation

```
CTACCTGAACGACGGCCAGTGAATTCGTGCGGGGACAGGACGGCTAGTGCAATGTTGACC (0/44)
                       G C       A                          (1/44)
                           AT                               (1/44)
```

Guanine Incorporation

```
CTACCTGAACGACGGCCAGTGAATTCGGGCGGGGACAGGACGGCTAGTGCAATGTTGACC (1/44)
```

Figure 2.6: Mutagenic profile for hPolη's replication of 17/73-mer-dG^{AP}. Results from the SOSA are shown separately based on specific nucleotide incorporated by Dpo4 in the control template. Sequences corresponding to primers used for amplification are shown in lower-font while sequenced nucleotides are upper font and underlined. Individual base substitutions (blue), deletions (red), and additions (green) are shown. Cyan highlighted font indicates position of the lesion.

Bar graphs depicting the number and types of mutation at each site over all the sequencing spectra for each enzyme can be found in Appendix 2. These graphs are useful for comparing the total mutations generated by each Y-family DNA polymerase and for visualizing the distribution of these mutations within the sequencing region.

Discussion

The importance of the Y-family DNA polymerases has become apparent due to their

widespread ability to replicate damaged DNA. Moreover, of the sixteen DNA polymerases identified in humans, four have been classified into this low-fidelity and error-prone family. In our study, we analyzed the bypass capabilities of two human Y-family DNA polymerases as well as Dpo4, the most widely studied Y-family DNA polymerase. It has become evident that, even within the same family, polymerases possess structural attributes that render their functions unique. *In vivo,* different polymerases will be more likely than others to bypass a specific lesion. Thus, it is necessary to determine the bypass efficiencies and mutagenic profiles to surmise which polymerase would likely bypass a specific lesion *in vivo.* As previously described, the mutagenic profile for each polymerase was determined using a novel experiment, known as SOSA. The efficiency of bypass and mutagenic profiles of hPolη and hΔPolκ were studied to determine which of these human DNA polymerase are more likely to bypass the dG^{AP} lesion. In addition, this data for the human Y-family DNA polymerases was compared to that of the model polymerase Dpo4, which is the sole Y-family DNA polymerase in *Sulfolobus sulfataricus*. In future projects, the human Y-family DNA polymerases hpolι and hRev1 will be investigated to compare their bypass efficiencies and mutagenic characterization with those of the polymerases used in this study.

Due to the interesting results from Chapter 1 as well as the predominance of 1-nitropyrene as an environmental pollutant, we decided to determine the bypass ability and characterize the mutagenic profiles of Dpo4 and two human Y-family DNA polymerases with the dG^{AP} lesion. The results of the running start assay from Chapter 1 showed that despite a decreased rate in the formation of full-length product, Dpo4 was able to bypass the dG^{AP} lesion. Moreover, we demonstrated that both hPolη and hΔPolκ were capable of bypassing this lesion and generate full-length product. Interestingly, the running start assays for Dpo4 and hΔPolκ

revealed the same two strong pause sites, indicating that each of these enzymes stalled in the incorporating a nucleotide directly across from the lesion and at the next downstream position. Moreover, these enzymes had a similarly slowed rate after encountering the dG^{AP} lesion. In sharp contrast, the running start assay for hPolη revealed no distinct pausing when encountering the lesion and was only mildly perturbed in the rate of production of full-length product. As Dpo4 is the only Y-family DNA polymerase in *Sulfalobus sulfataricus*, it is clearly the enzyme that would bypass the dG^{AP} lesion *in vivo*. However, in humans, one of the four Y-family DNA polymerases would most likely bypass this specific lesion. Certainly, when comparing only the running start assays of hPolη and hΔPolκ, we would surmise that hPolη would more likely bypass the dG^{AP} lesion *in vivo*. This conclusion is based solely on hPolη's lack of the significant stalling demonstrated by hΔPolκ and due to its rate of formation of full-length product in comparison to hΔPolκ.

However, we collected more data regarding the bypass of the dG^{AP} lesion that provided us with the precise DNA sequence synthesized by each DNA polymerase. This data was collected via the implementation of SOSA, a novel assay developed in the Suo laboratory. The SOSA experiment was tested by sequencing the replication of control 73-mer by Dpo4. The sequencing results revealed only four sequences with errors, two were substitution and two were deletions. As previously stated, this resulted in an error frequency that corresponded to the fidelity determined by pre-steady state kinetic assays in Chapter 1 (10^{-3}-10^{-5}). Thus, we confirmed that SOSA was a valid way of sequencing the bypass products and determined that there was negligible sequence bias.

As described in the Results section, Dpo4 preferentially incorporated the correct

nucleotide, while hΔPolκ and hPolη tended to synthesize across from the lesion more indiscriminately. A comparison of these three Y-family DNA polymerases and their DNA synthesis opposite the dG^{AP} lesion is found in Figure 2.7. Figure 2.7 makes it very clear how much more efficient Dpo4 is in its synthesis opposite the lesion in comparison to the human Y-family DNA polymerases. In comparing hΔPolκ and hPolη and their incorporation at the site of the lesion, they appear to have very similar tendencies to insert mutations rather than the correct nucleotide.

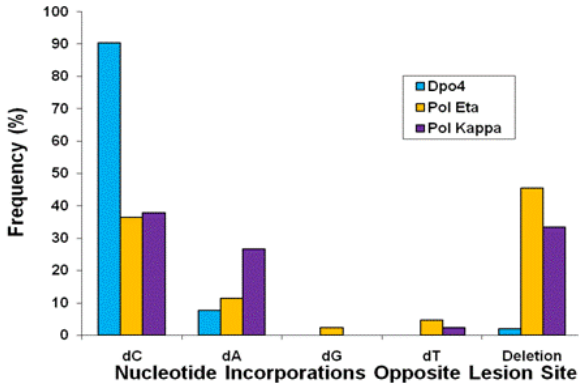


Figure 2.7: Comparison of the synthesis opposite the lesion for Dpo4 (blue), hPolη (yellow), and hΔPolκ (purple).

Besides the incorporation opposite the lesion, the mutagenic profile of the extension of the bypass product directly one position downstream of the lesion was also analyzed (Figure 2.8). At this position, the three Y-family DNA polymerases are more likely to incorporate the correct nucleotide. At this site, Dpo4 continues to generate the most accurate bypass products. When comparing hΔPolκ and hPolη, hΔPolκ has a greater fidelity at this position than hPolη. This decreased fidelity of hPolη is due to its increased tendency to generate a deletion at this position.

Finally, it is important to analyze and compare the overall fidelity of each polymerase.

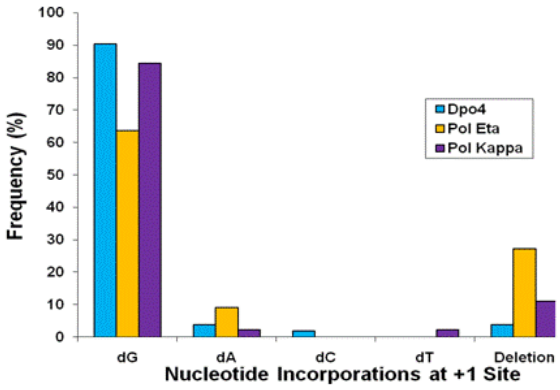


Figure 2.8: Comparison of the synthesis of the first downstream position after the dG^{AP} lesion for Dpo4 (blue), hPolη (yellow), and hΔPolκ (purple).

Figures A2.1, A2.2, and A2.3 in the appendix depict the mutations at each position in the sequencing window for Dpo4, hΔPolκ, and hPolη, respectively . Quantitatively, Dpo4 has an overall fidelity of 0.01, while the fidelities of hΔPolκ and hPolη respectively are 0.06 and .11. Based on this study, the presence of the dG^{AP} lesion decreases the fidelity of Dpo4 by only about ten fold. Moreover, while hPolη seemed to have the greatest bypass efficiency based on the running start assay, its mutagenic profile is significantly worse than the other human Y-family DNA polymerase analyzed.

Dpo4 is the only Y-family DNA polymerase in *Sulfolobus sulfataricus* and thus will be responsible for translesion synthesis across the dG^{AP} lesion. Fortunately, for this organism, it seems from our results that Dpo4 bypasses the lesion with comparatively good fidelity. Unlike *Sulfolobus sulfataricus,* humans have four Y-family DNA polymerases. The evolution of

numerous enzymes in the same organism may be to decrease the errors associated in this process by having polymerases that specialize in the bypass of a particular lesion. This theory is supported by the fact that human Y-family DNA polymerases possess drastically different abilities to bypass different DNA lesions. Thus, the results of this study can be used in an attempt to understand the precise mutagenic characterization of two human Y-family DNA polymerases in response to a specific lesion. Despite its increased bypass efficiency, hPolη generated significantly more mutations than hΔPolκ. Thus, based solely on the comparison of the mutagenic profiles it is more likely that hΔPolκ bypasses the dG^{AP} lesion *in vivo*. However, more studies must be done to analyze the other human Y-family DNA polymerases. The method used in this study can be used in future studies to generate information regarding the mutagenic characterization of each Y-family DNA polymerase for a variety of DNA lesions.

Chapter 3: Antiviral Drugs as Substrates for X- and Y-family DNA Polymerases

Background and Introduction

In modern society, viral infection is one of the leading causes of disease. While there are a plethora of different virally induced diseases, two stand at the forefront due to their widespread effect on human health. Hepatitis B virus (HBV) causes severe liver disease and chronically inflicts 350 million people worldwide [43]. Human Immunodeficiency virus (HIV) attacks the human immune system, depriving the body of fighting off infection. Currently, 33 million people are living with this dangerous disease[44] . These viral infections are very aggressive due to the precision of the virus life cycle, especially replication within the host.

Due to the potency of these diseases and their widespread effect on human health, many drugs have been developed to treat these viral infections. While there are many kinds of treatments used to allay these infections, this study will focus on a particular class of drugs, known as nucleoside analogs. Nucleoside analogs are a class of antiviral drugs that mimic natural nucleotides due to their similar structure and inhibit viral replication in infected host cells. These drugs are administered as nucleosides, however, they are phosphorylated by the kinases of the host cell [45]. Currently twelve different nucleoside analog drugs have been approved by the Food and Drug Adminstration (FDA) for the treatment of HBV and HIV alone. Of these twelve drugs, this study will focus on five nucleoside analogs. A structural representation of these drugs can be seen in Figure 3.1.

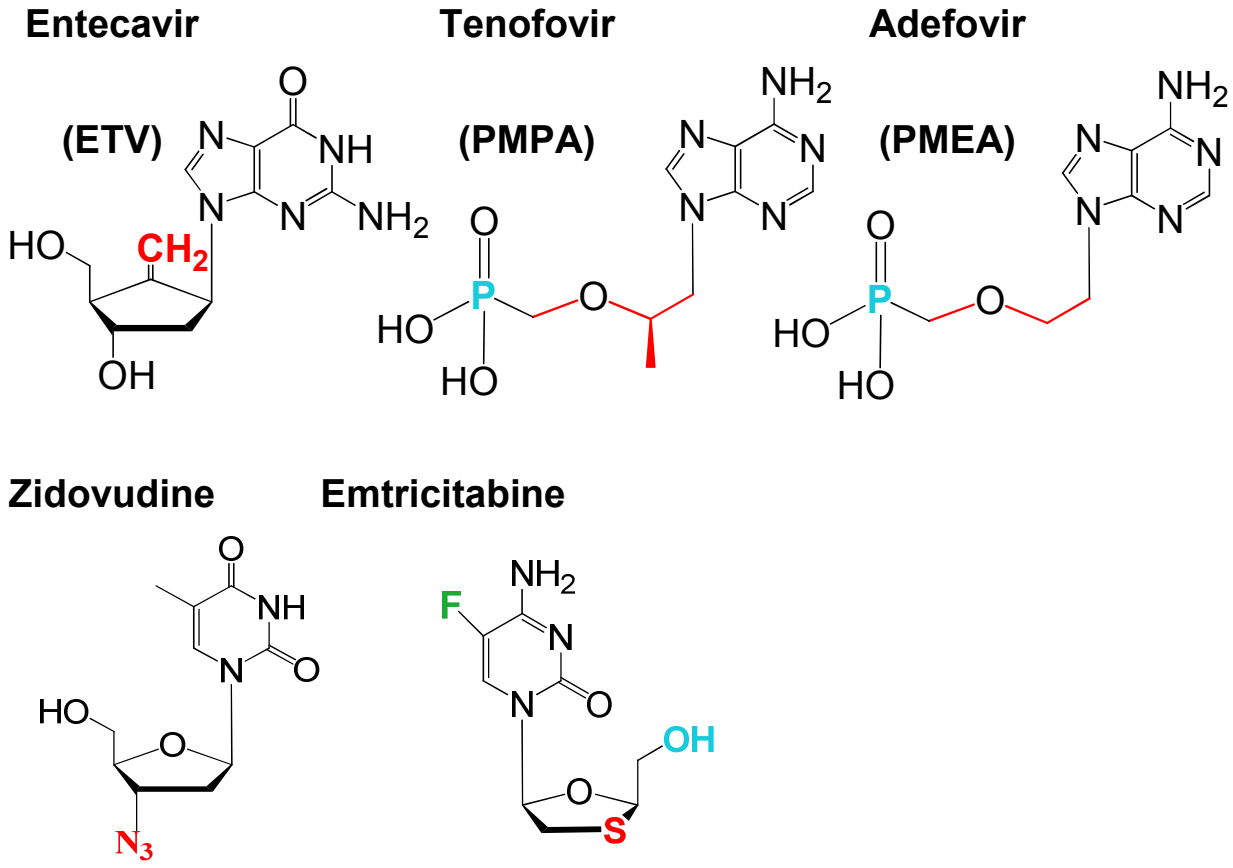


Figure 3.1: Structure of nucleoside analogs used for this study. Red represents changes to the ribose. Green represents modifications to the base. Blue represents other types of modifications

Of these drugs, Entecavir (ETV) and Adefovir (PMEA) are used to treat HBV [46], while Zidovudine (AZT) and Emtricitabine (L-FTC) are used for treatment of HIV [47]. Tenofovir (PMPA) has been approved to treat both HBV and HIV. These analogs are different from the natural nucleotides due to modifications to the ribose sugar, change to the base, or inversion of orientation. One important feature of many of these drugs is that they lack a 3' hydroxyl group.

Thus, if these drugs are incorporated, replication will be terminated.

When administered, there is nothing about the structure of these drugs that inhibits their incorporation into the human genome by the host DNA polymerases. If these drugs are incorporated into the host genome, they will discontinue replication of the genome, which will lead to cell death. Thus, these drugs may be substrates for the human DNA polymerases and incorporation into the human genome of these nucleotide analogs could be the cause of many of the side effects associated with these drugs [48]. However, before these nucleoside analogs are approved by the FDA, they are tested with the "classical" DNA polymerases to ensure a low selectivity of the nucleotide analog over the natural nucleotide. Thus, many studies have shown that the "classical" DNA polymerases have a low probability of incorporating these drugs. Conversely, no studies have been done investigating these analogs with the recently discovered X- and Y- family DNA Polymerases. Thus, the aim of this project will be to investigate if these nucleoside analogs are substrates for X- and Y-family DNA polymerases.

The human X-family DNA polymerases include human DNA polymerase Lambda (hPolλ), Beta (hPolβ), mew (hPolμ), and terminal deoxynucleotidyl transferase (TdT). Despite belonging to the same family, these polymerases have quite different *in vivo* functions. For example, TdT randomly incorporates nucleotides at the ends of DNA during V(D)J recombination, while hPolβ is a major participant in Base Excision Repair (BER) [49]. Overall, these X-family DNA polymerases tend to play a role in several DNA repair pathways. hPolλ was selected as the model X-family DNA polymerase to determine if the aforementioned nucleotide analogs would be substrates for this essential family of enzymes. hPolλ was chosen due to its overall sequence homology to the other members of its family. In addition, the role of hPolλ is

for the most part unknown. However, hPolλ has been suggested to have a role in several DNA repair pathways including DNA repair synthesis in meiosis [50], "short patch" BER [51], in the proliferating cell nuclear antigen-dependent BER [52], and in the repair of double stranded breaks through base excision repair [53].

Besides determining if nucleoside analogs are substrates for the repair enzymes of the X-family, we also analyzed whether these drugs would be incorporated by the Y-family of DNA polymerases. The Y-family DNA polymerases have been mentioned in the previous chapters due to their role in translesion synthesis. As mentioned previously, four Y-family DNA polymerases have been identified in humans. These include human DNA polymerases eta (hPolη), iota (hPolι), kappa (hPolκ), and Rev1. hPolη was chosen as the model Y-family DNA polymerase due to its known function in cis-syn thymine-thymine dimer lesion bypass [27]. Thus, we determined whether five nucleotide analogs would be substrates for model X-and Y-family DNA polymerases, hPolλ and hPolη, respectively.

Methods

Materials- For this project, OptiKinase from United Biochemical (Cleveland, OH), [γ-^{32}P]ATP from GE Healthcare, and dNTPs from Invitrogen were purchased. The full length hPolλ and hPol η were expressed in *E. coli* and purified within the Suo laboratory.

Synthetic Oligonucleotides- DNA templates and primers (Table 3.1) were purchased from Integrated DNA Technologies (Coralville, IA) and purified using denaturing PAGE. The concentration of these DNA oligonucleotides was determined using UV spectroscopy at 260 nm.

Table 3.1 Sequences of DNA oligonucleotides

Primers	
21-mer	5′-CGCAGCCGTCCAACCAACTCA-3′
19-mer-5'P	5′-PCGTCGATCCAATGCCGTCC-3′
19-mer2-5'P	5′-PAGTCGATCCAATGCCGTCC-3′
Templates	
41-mer D-8	3′-GCGTCGGCAGGTTGGTTGAGT**C**GCAGCTAGGTTACGGCAGG-5′
41-mer D-6t	3′-GCGTCGGCAGGTTGGTTGAGT**G**TCAGCTAGGTTACGGCAGG-5′
41-mer D-1	3′-GCGTCGGCAGGTTGGTTGAGT**A**GCAGCTAGGTTACGGCAGG-5′
41-mer D-7	3′-GCGTCGGCAGGTTGGTTGAGT**T**GCAGCTAGGTTACGGCAGG-5′

Radiolabeling and Annealing DNA substrates- The 21-mer primer was 5'-^{32}P-labeled by incubating it with Optikinase and [γ-^{32}P]ATP for 3 hours at 37 ˚C. This radiolabeled primer was annealed to an unlabeled template, in which the nucleotide analog would be a "correct" incorporation, at a molar ratio of 1.00:1.15. Due to the fact that hPolλ likely has a role in DNA gap filling, templates for hPolλ experiments were also annealed to a downstream primer, 19-mer-5'P or 19mer2-5'P. The annealing mixtures were heated to 95 ˚C for five minutes and then slowly cooled to room temperature.

Buffers: Pre-steady-state kinetic assays for hPolη were performed in optimized buffer R (50 mM HEPES, pH 7.5 at 37 ˚C, 5 mM $MgCl_2$, 50 mM NaCl, 0.1 mM EDTA, 5mM dithiothreitol, 10% glycerol, and 0.1 mg/mL bovine serum albumin). For hPolλ, pre-steady-state kinetic assays were performed in buffer L (50mM Tris-Cl(pH 8.4 at 37 °C), 5 mM $MgCl_2$, 100 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, 0.1 mg/mL BSA). The concentrations indicated above were those of the final mixed solutions.

Determination of Selectivity Factor- Like in Chapter 1, pre-steady state single turnover

dNTP incorporation assays were employed to determine the k_p and $K_{d,\,dNTP}$, where the k_p is the maximum rate of incorporation of a dNTP and $K_{d,\,dNTP}$ is the equilibrium dissociation constant of an incoming nucleotide. In brief, a solution of 5'-^{32}P-labeled DNA (30 nM) and Dpo4 (120 nM) in buffer R or L was pre-incubated at 37 ˚C and subsequently mixed with solutions containing increasing amounts of a single natural nucleotide or nucleotide analog. These reactions were quenched using the rapid-chemical quench or manually after various time points using 0.37 M EDTA. The formed product was separated from the initial primer by denaturing PAGE (17% acrylamide, 8 M urea) and quantitated with a Phosphorimager 445 SI. For each concentration of natural nucleotide or corresponding analog, the function of product formation per time point was fit to a single exponential equation (Equation 1.2). Where k_{obs} is the observed reaction rate constant, and A is the reaction's product formation amplitude. Next, the values of the k_{obs} were plotted against the concentration of dNTP for the respective reaction. This relationship was fit to a hyperbolic equation (Equation 1.3). Where, as stated previously, k_p is the maximum rate of dNTP incorporation, and the $K_{d,dNTP}$ is the dissociation constant for the ternary complex (Dpo4·DNA·dNTP) at equilibrium. Combining these terms, the substrate specificity was determined by taking the ratio of the k_p over the $K_{d,dNTP}$. The selectivity factor was determined by comparing the substrate specificity of the natural nucleotide over the substrate specificity of the nucleotide analog. Here, a higher selectivity factor would indicate a high probability of incorporation of the natural nucleotide over the nucleotide analog drug.

Results

Determining the Selectivity Factors for each Nucleotide Analog- As stated earlier, the selectivity factor is determined by comparing the ratio of the substrate specificities of the natural nucleotide to the nucleotide analog. The substrate specificity is a term that combines the k_p and $K_{d,DNA}$ by taking the former over the later (k_p/ $K_{d,DNA}$). In order to determine the k_p and $K_{d,DNA}$, we performed single dNTP incorporation assays under single-turnover reaction conditions. In this assay, a pre-incubated solution of DNA polymerase (either hPolλ or hPolη) (120 nM) and 5'-^{32}P-labeled DNA substrate are rapidly mixed with either the natural nucleotide or nucleotide analog and quenched at various time points with .37M EDTA. The products are then separated using PAGE. The product formation is then plotted against time and fit to Equation 1.2 to determine the observed reaction rate (k_{obs}) (Figure3.2).

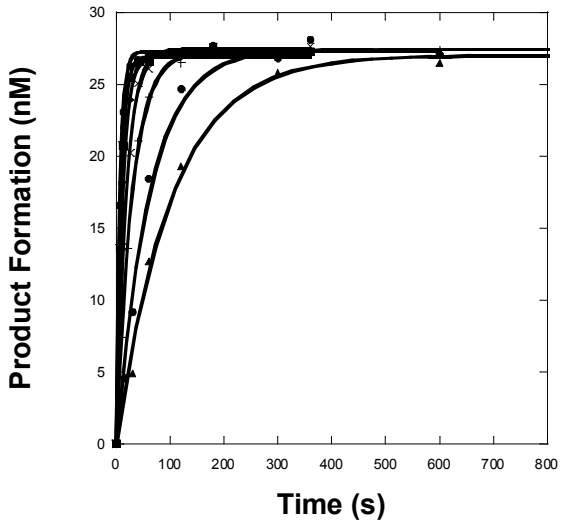


Figure 3.2: Relationship of product formation versus time for PMPA and hPolη. Each time course was fit to Eq 1.2 to yield a k_{obs}.

The dependence of k_{obs} on the concentration of the dNTP is then plotted and fit to Equation 1.3, which yields values for the k_p and $K_{d,\ DNA}$ (Figure3.3). The determined kinetic results of the incorporation of both natural nucleotides and nucleotide analogs by hPolλ and hPolη are summarized in Table 3.2 and Table 3.3.

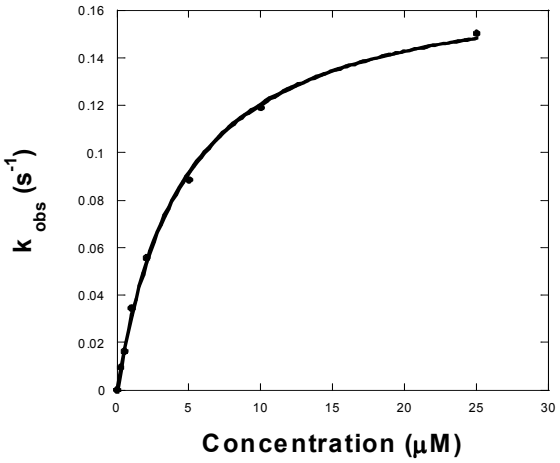


Figure 3.3: The Plot of Observed Rate Constant against Concentration. The plot of k_{obs} values against PMPA concentrations was fit to Eq 1.3.

Table 3.2: Kinetic Parameters for hPolλ's incorporation of a natural nucleotide or nucleotide analog

NA	k_p (s^{-1})	K_d (μM)	k_p/K_d ($\mu M^{-1}s^{-1}$)	Selection Factor[a]
dCTP	1.57 ± 0.04	0.9 ± 0.1	1.7	
L-FTC-TP	0.0049 ± 0.0001	0.36 ± 0.03	0.014	121
dATP	1.5 ± 0.1	0.9 ± 0.3	1.7	
PMPA-DP	0.095 ± 0.008	3.7 ± 0.9	0.026	65
PMEA-DP	0.175 ± 0.004	4.5 ± 0.3	0.039	44
dGTP	2.5 ± 0.1	2.1 ± 0.3	1.2	
ETV-TP	0.034 ± 0.002	0.09 ± 0.02	0.38	3.2
dTTP	3.9 ± 0.2	2.6 ± 0.4	1.5	
AZT-TP	1.83 ± 0.09	12 ± 2	0.15	10

[a]Calculated as $(k_p/K_d)_{dNTP}/(k_p/K_d)_{NA}$.

Table 3.3: Kinetic Parameters for hPolη's incorporation of a natural nucleotide or nucleotide analog

NA	k_p (s^{-1})	K_d (μM)	k_p/K_d ($\mu M^{-1}s^{-1}$)	Selection Factor[a]
dCTP	49 ± 2	25 ± 4	1.96	
L-FTC-TP	0.0244 ± 0.0004	2.7 ± 0.1	9.0×10^{-3}	217
dATP	35 ± 3	130 ± 26	0.27	
PMPA-DP	0.0134 ± 0.007	90 ± 10	1.5×10^{-4}	1,800
PMEA-DP	0.069 ± 0.008	55 ± 18	1.3×10^{-3}	208
dGTP	38 ± 2	80 ± 12	0.48	
ETV-TP	0.24 ± 0.01	2.4 ± 0.4	0.1	5
dTTP	35 ± 1	41 ± 5	0.85	
AZT-TP	0.31 ± 0.05	478 ± 155	6.5×10^{-4}	1,300

[a]Calculated as $(k_p/K_d)_{dNTP}/(k_p/K_d)_{NA}$.

By dividing the k_p by the $K_{d,\,DNA}$, these terms were combined to determine the substrate specificity. To analyze the potential of each nucleotide analog to be incorporated by either DNA polymerase, the substrate specificity of the natural nucleotide was compared to that of the nucleotide analog. Thus, the selectivity factor was determined by taking the ratio of the substrate specificity of the natural nucleotide over substrate specificity of the nucleotide analog.

The pre-steady state kinetic analysis of hPolλ and hPolη's incorporation of the various nucleotide analogs indicated that these drugs could be incorporated into the human genome. Interstingly, the hPolλ and hPolη had different order of preferences for the studied group of nucleotide analogs. hPolλ's substrate preference was L-FTC-TP< PMPA-DP < PMEA-DP<AZT-TP< ETV-TP. The substrate preference of hPolη was PMPA-DP<AZT-TP<L-FTC-

TP<PMEA-DP<ETV-TP. Thus, while both hPolλ and hPolη are capable of incorporating these nucleotide analogs, they do so with different specificities.

Discussion

Nucleoside analog drugs play a significant role in the treatment of viral infections, especially HBV and HIV. They function to inhibit the viral replication machinery and reduce the rate of replication of the virus within the host. These drugs, however, may also interfere with the human host's DNA replication. While some research has been done studying the "classical" DNA polymerases ability to select the natural nucleotide over the analog, currently, no studies have been performed to determine the selectivity of other families of polymerases. Thus, this project investigated the selectivity of incorporation for polymerases that model the X- and Y-family DNA polymerases, hPolλ and hPolη, respectively.

Through the utilization of pre-steady state kinetic assays, we were able to determine the microscopic kinetic parameters for the incorporation of each nucleotide analog and corresponding natural nucleotide by either hPolλ or hPolη. We compared the substrate specificites for each nucleotide analog with the corresponding natural nucleotide to determine the selectivity of the polymerase for the natural nucleotide over the analog. Thus, we used these selectivity factors to demonstrate the likelihood of each nucleotide analog to be incorporated by the polymerase.

A high selectivity factor is associated with a low probability of incorporation of the nucleotide analog. When a drug is being developed, the company tries to identify analogs with a

low selectivity for the viral replication machinery but a high selectivity for the human DNA polymerases. In addition, the pharmaceutical companies follow a standard that the selectivity factor for incorporation of the analog by the "classical" replicative polymerases must be greater than 500, indicating that the drug is a poor substrate for these enzymes. Comparing this standard to the results in Tables 3.2 and 3.3, it is obvious that many of these drugs would not meet the standard if it was broadened to the DNA polymerases responsible for translesion synthesis and DNA repair. All the nucleoside analog drugs that were tested with hPolλ had a selectivity factor less than the standard (Figure 3.4).

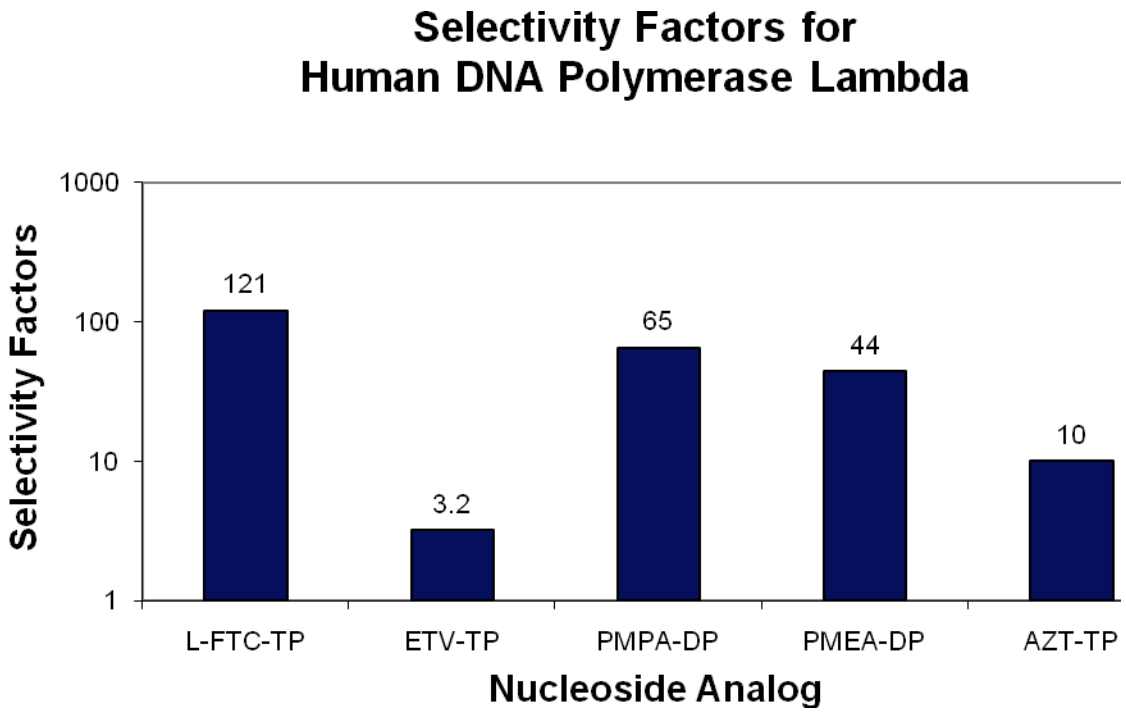


Figure 3.4: Selectivity factors for hPolλ

Thus, all these drugs may be substrates for the X-family DNA polymerases *in vivo*. hPolη had

selectivity factors lower than the standard for ETV, PMEA, and L-FTC. Based on these results, these drugs would likely be substrates for Y-family DNA polymerases. Contrastingly, the selectivity factors for hPolη's incorporation of PMPA and AZT were well above the 500 standard. Thus, these drugs would be unlikely to be incorporated by the Y-family. A comparison of the selectivity factors for the aforementioned nucleotide analogs is shown in Figure 3.5.

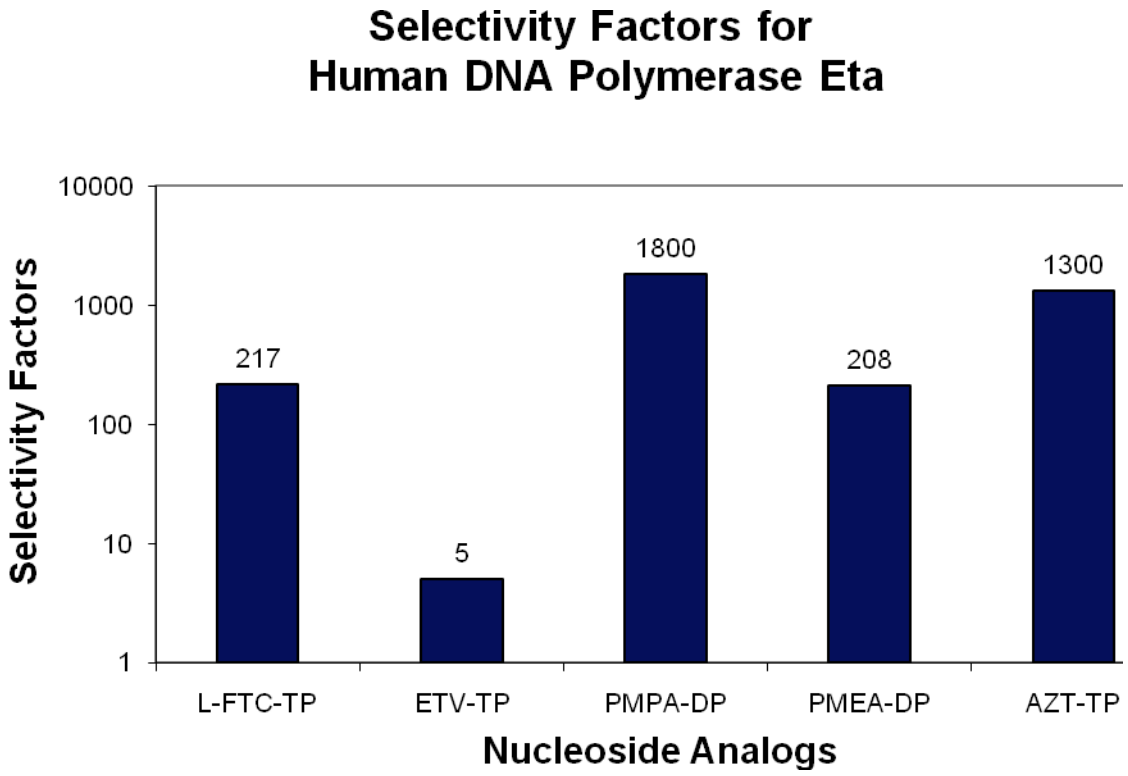


Figure 3.5: Selectivity factors for hPolη

Interestingly, the selectivity factors for ETV's incorporation by hPolλ and hPolη were the lowest of all the other drugs tested for both enzymes. Moreover, the selectivity factors for ETV incorporation were less than 10. This data indicates that the incorporation of ETV into the

human genome by both X- and Y-family DNA polymerases is almost as probable as incorporation of a natural nucleotide. The reason for these low selectivity factor seems to arise from the tight binding of ETV to the ternary complex rather than an increased rate of incorporation.

The low selectivity factors of incorporation of ETV as well as many of the other nucleotide analogs by hPolλ and hPolη indicate that these drugs could be incorporated into the human genome. These polymerases seem unable to distinguish the analogs from their corresponding natural nucleotides. If the X- and Y- family DNA polymerases are employed and a patient is being treated with one of the aforementioned nucleoside analogs, these enzymes will likely incorporate the analog and thus terminate replication of the human genome. The results presented could provide insight into the toxicity and side effects associated with treatment consisting of nucleoside analog drugs. More specifically, he observed clinical toxicity may be related to the potential incorporation of these nucleotide analogs catalyzed by human X- and Y-family DNA polymerases *in vivo*.

Conclusion

In Chapter 1, the bypass mechanism of a single base lesion, dG^{AP}, was determined for a model Y-family DNA polymerase, Dpo4. We found that while Dpo4 is capable of bypassing the dG^{AP} lesion, its rate is significantly perturbed when encountering the lesion. Moreover, the polymerase stalls in its incorporation directly across from lesion and in the extension of the bypass product one nucleotide downstream. This stalling was largely due to the decrease in correct dNTP incorporation efficiencies at the pause sites. Finally, we learned that the bypass of the dG^{AP} lesion demonstrated biphasic kinetics. From this finding, we hypothesized that the some of the ternary complexes were bound in an unproductive conformation and thus required transformation to a more productive complex.

In Chapter 2, the mutagenic profiles for the bypass of the dG^{AP} lesion by Dpo4, hΔPolκ, and hPolη were determined using a novel assay, known as SOSA. We utilized this assay to sequence the bypass products of the replication opposite a template containing dG^{AP}. The sequencing results from this chapter showed the individual substitutions, additions, and deletions that occurred in bypass of the dG^{AP} lesion. Our study showed that Dpo4 sequenced a correct dCTP opposite the lesion in 90% of the sequence spectra. Contrastingly, hΔPolκ, and hPolη indiscriminately synthesized the bypass product opposite the lesion. When comparing, these two human Y-family DNA polymerases, it seems that although hPolη bypasses the lesion with more kinetic efficiency, hΔPolκ maintains a greater fidelity.

Finally, in Chapter 3, we analyzed the potential of nucleoside analog drugs to be

incorporated by X- and Y-family DNA polymerases, which were modeled by hPolλ and hPolη, respectively. Nucleoside analogs are meant to disrupt the replication machinery of the virus to slow its growth within the host. These analogs may become substrates for the human DNA polymerases, causing the replication of the human genome to be terminated. The microscopic kinetic parameters for the incorporation of both nucleotide analogs and their corresponding natural nucleotide were determined using pre-steady state kinetic assays. These parameters were compared to determine the selectivity factor for each nucleoside analog. The results showed that many of the analogs studied had low selectivity, indicating that they would likely be incorporated by either the X- or Y-family DNA polymerases. The demonstrated potential of these nucleoside analogs to be incorporated into the human genome by X- and Y-family DNA polymerases may relate to the observed clinical toxicity of many of these drugs.

Overall, this thesis outlines my progression as a researcher in the Suo laboratory. It provides detailed explanations of a few of my projects in the lab that have mainly been in collaboration with Jessica Brown and Shanen Sherrer. Besides, contributing to my growth as a researcher, these projects have contributed to general scientific understanding of X- and Y-family DNA polymerases.

Appendix 1

Supplementary Table 1. Kinetic parameters of dNTP incorporation into undamaged DNA (26-mer)

	$K_{d,\ dNTP}$ (μM)	k_p (s^{-1})	$k_p/K_{d,\ dNTP}$ ($\mu M^{-1}s^{-1}$)	Fidelity[a]
19/26-mer				
dGTP	183 ± 54	4.6 ± 0.4	2.5×10^{-2}	-
dATP	731 ± 175	$(1.0 \pm 0.1) \times 10^{-2}$	1.4×10^{-5}	5.6×10^{-4}
dCTP	350 ± 107	$(5.0 \pm 0.5) \times 10^{-2}$	1.4×10^{-4}	5.6×10^{-3}
dTTP	1440 ± 305	$(1.3 \pm 0.2) \times 10^{-2}$	8.9×10^{-6}	3.6×10^{-4}
20/26-mer				
dCTP	205 ± 64	11.6 ± 0.9	5.7×10^{-2}	-
dATP	631 ± 136	$(1.2 \pm 0.1) \times 10^{-2}$	1.9×10^{-5}	3.3×10^{-4}
dGTP	77 ± 16	$(3.5 \pm 0.2) \times 10^{-3}$	4.5×10^{-5}	7.9×10^{-4}
dTTP	489 ± 52	$(4.0 \pm 0.2) \times 10^{-2}$	8.3×10^{-5}	1.5×10^{-3}
21/26-mer				
dGTP	437 ± 19	1.62 ± 0.03	3.7×10^{-3}	-
dATP	859 ± 168	$(2.0 \pm 0.2) \times 10^{-3}$	2.4×10^{-6}	6.5×10^{-4}
dCTP	701 ± 30	$(2.14 \pm 0.04) \times 10^{-2}$	3.1×10^{-5}	8.3×10^{-3}
dTTP	1180 ± 144	$(2.1 \pm 0.2) \times 10^{-3}$	1.8×10^{-6}	4.9×10^{-4}
22/26-mer				
dCTP	129 ± 19	4.4 ± 0.2	3.4×10^{-2}	-
dATP	1313 ± 154	$(5.5 \pm 0.4) \times 10^{-3}$	4.2×10^{-6}	1.2×10^{-4}
dGTP	567 ± 95	$(4.8 \pm 0.4) \times 10^{-3}$	8.5×10^{-6}	2.5×10^{-4}
dTTP	1340 ± 454	$(1.5 \pm 0.3) \times 10^{-2}$	1.1×10^{-5}	3.2×10^{-4}
23/26-mer				
dGTP	116 ± 24	2.8 ± 0.1	2.4×10^{-2}	-
dATP	431 ± 25	$(6.6 \pm 0.1) \times 10^{-3}$	1.5×10^{-5}	6.2×10^{-4}
dCTP	918 ± 102	$(1.5 \pm 0.1) \times 10^{-3}$	1.6×10^{-6}	6.7×10^{-5}
dTTP	1220 ± 59	$(3.8 \pm 0.1) \times 10^{-3}$	3.1×10^{-6}	1.3×10^{-4}

[a]Calculated as $(k_p/K_{d,\ dNTP})_{incorrect}/[(k_p/K_{d,\ dNTP})_{correct} + (k_p/K_{d,\ dNTP})_{incorrect}]$.

Appendix 2

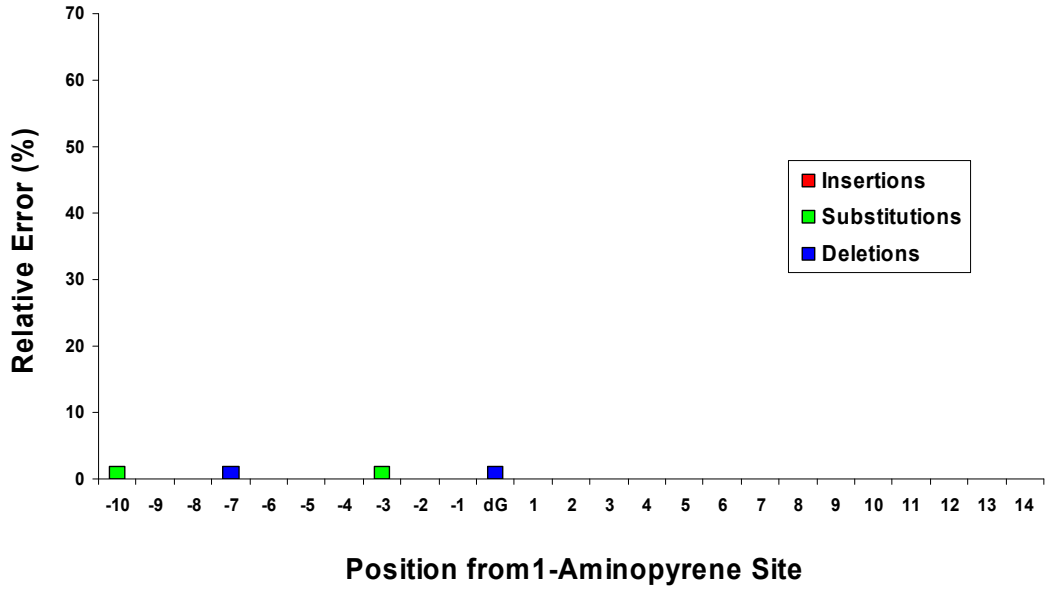


Figure A2.1: Mutations at each position for Dpo4's replication of 17/73-mer.

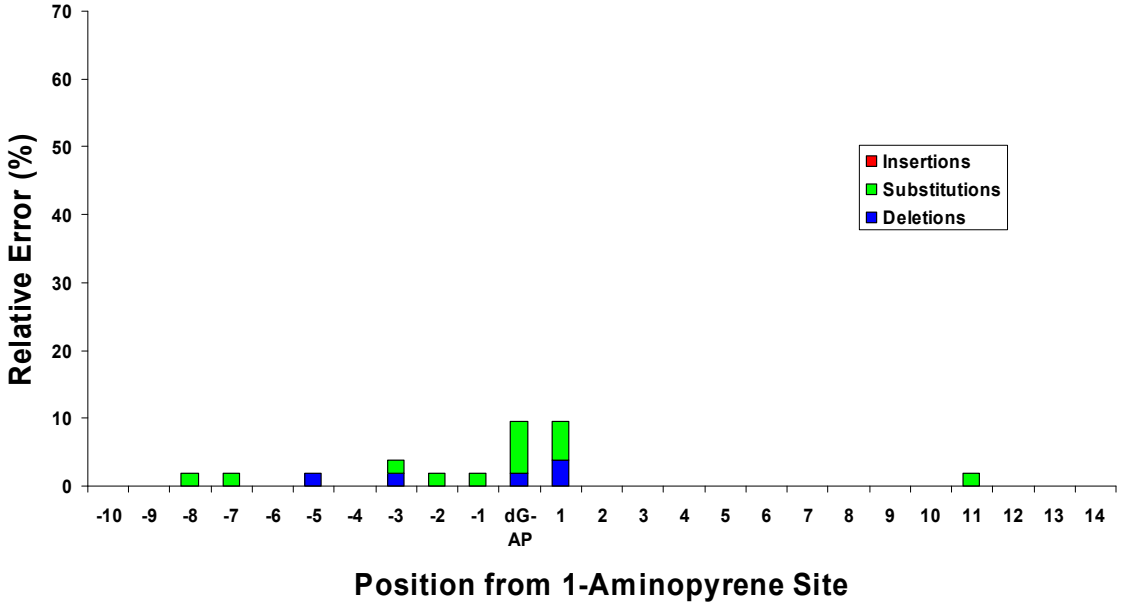


Figure A2.2: Mutations at each position for Dpo4's replication of 17/73-mer-dG^{AP}.

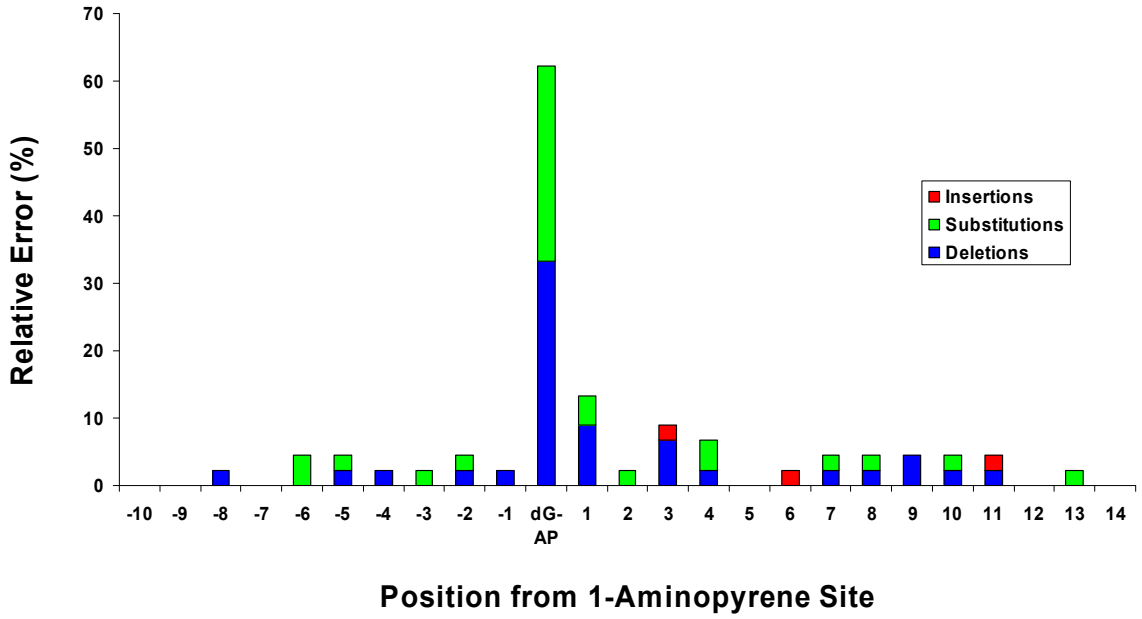


Figure A2.3: Mutations at each position for hΔPolκ's replication of 17/73-mer-dG^{AP}.

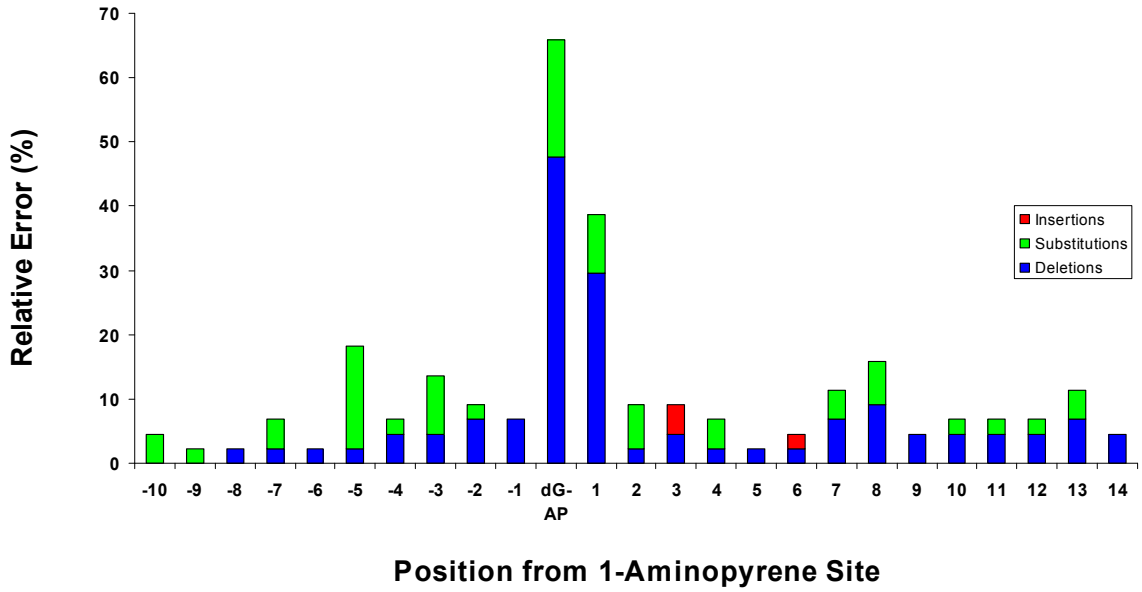


Figure A2.4: Mutations at each position for hPolη's replication of 17/73-mer-dG^{AP}.

References

1. Sherrer, S.M., et al., *Mechanistic studies of the bypass of a bulky single-base lesion catalyzed by a Y-family DNA polymerase.* J Biol Chem, 2009. **284**(10): p. 6379-88.
2. Pohjola, S.K., et al., *DNA binding of polycyclic aromatic hydrocarbons in a human bronchial epithelial cell line treated with diesel and gasoline particulate extracts and benzo[a]pyrene.* Mutagenesis, 2003. **18**(5): p. 429-38.
3. Pohjola, S.K., et al., *Polycyclic Aromatic Hydrocarbons of Diesel and Gasoline Exhaust and DNA Adduct Detection in Calf Thymus DNA and Lymphocyte DNA of Workers Exposed to Diesel Exhaust.* Polycyclic Aromatic Compounds, 2004. **24**(4-5): p. 451-465.
4. Liu, Y., et al., *Interactions of human replication protein A with single-stranded DNA adducts.* Biochem. J., 2005. **385**(Pt 2): p. 519-26.
5. Zou, Y., et al., *Effects of DNA adduct structure and sequence context on strand opening of repair intermediates and incision by UvrABC nuclease.* Biochemistry, 2003. **42**(43): p. 12654-61.
6. Mizukami, S., et al., *Varying DNA base-pair size in subangstrom increments: evidence for a loose, not large, active site in low-fidelity Dpo4 polymerase.* Biochemistry, 2006. **45**(9): p. 2772-8.
7. Fiala, K.A. and Z. Suo, *Pre-Steady-State Kinetic Studies of the Fidelity of Sulfolobus solfataricus P2 DNA Polymerase IV.* Biochemistry, 2004. **43**(7): p. 2106-15.
8. Fiala, K.A., C.D. Hypes, and Z. Suo, *Mechanism of abasic lesion bypass catalyzed by a Y-family DNA polymerase.* J. Biol. Chem., 2007. **282**(11): p. 8188-98.
9. Johnson, R.E., L. Prakash, and S. Prakash, *Distinct mechanisms of cis-syn thymine dimer bypass by Dpo4 and DNA polymerase eta.* Proc. Natl. Acad. Sci. U. S. A., 2005. **102**(35): p. 12359-64.
10. Brown, J.A., et al., *Mechanism of double-base lesion bypass catalyzed by a Y-family DNA polymerase.* Nucleic Acids Res., 2008. **36**(12): p. 3867-78.
11. Bauer, J., et al., *A structural gap in Dpo4 supports mutagenic bypass of a major benzo[a]pyrene dG adduct in DNA through template misalignment.* Proc. Natl. Acad. Sci. U. S. A., 2007. **104**(38): p. 14905-10.
12. Sabbioni, G. and C.R. Jones, *Biomonitoring of arylamines and nitroarenes.* Biomarkers, 2002. **7**(5): p. 347-421.
13. Mitchelmore, C.L., D.R. Livingstone, and J.K. Chipman, *Conversion of 1-nitropyrene by Brown trout (Salmo trutta) and turbot (Scophtalamus maximus) to DNA adducts detected by 32P-postlabelling.* Biomarkers, 1998. **3**(1): p. 21 - 33.
14. Malia, S.A., R.R. Vyas, and A.K. Basu, *Site-specific frame-shift mutagenesis by the 1-nitropyrene-DNA adduct N-(deoxyguanosin-8-y1)-1-aminopyrene located in the (CG)3 sequence: effects of SOS, proofreading, and mismatch repair.* Biochemistry, 1996. **35**(14): p. 4568-77.
15. Fiala, K.A. and Z. Suo, *Mechanism of DNA Polymerization Catalyzed by Sulfolobus solfataricus P2 DNA Polymerase IV.* Biochemistry, 2004. **43**(7): p. 2116-25.
16. Zang, H., et al., *DNA adduct bypass polymerization by Sulfolobus solfataricus DNA polymerase Dpo4: analysis and crystal structures of multiple base pair substitution and frameshift products with the adduct 1,N2-ethenoguanine.* J. Biol. Chem., 2005. **280**(33): p. 29750-64.
17. Ling, H., et al., *Snapshots of replication through an abasic lesion; structural basis for base substitutions and frameshifts.* Mol. Cell, 2004. **13**(5): p. 751-62.
18. Ling, H., et al., *Replication of a cis-syn thymine dimer at atomic resolution.* Nature, 2003. **424**(6952): p. 1083-7.
19. She, Q., et al., *The complete genome of the crenarchaeon Sulfolobus solfataricus P2.* Proc. Natl. Acad. Sci. U. S. A., 2001. **98**(14): p. 7835-40.
20. Ohashi, E., et al., *Fidelity and processivity of DNA synthesis by DNA polymerase kappa, the*

product of the human DINB1 gene. J. Biol. Chem., 2000. **275**(50): p. 39678-84.
21. Wooster, R., et al., *Instability of short tandem repeats (microsatellites) in human cancers.* Nat. Genet., 1994. **6**(2): p. 152-6.
22. Nolan, S.J., et al., *Solution properties and computational analysis of an oligodeoxynucleotide containing N-(deoxyguanosin-8-yl)-1-aminopyrene.* Carcinogenesis, 1996. **17**(1): p. 133-44.
23. Fiala, K.A., et al., *Mechanism of Template-independent Nucleotide Incorporation Catalyzed by a Template-dependent DNA Polymerase.* J. Mol. Biol., 2007. **365**(3): p. 590-602.
24. Gu, Z., et al., *Solution structure of the N-(deoxyguanosin-8-yl)-1-aminopyrene ([AP]dG) adduct opposite dA in a DNA duplex.* Biochemistry, 1999. **38**(33): p. 10843-54.
25. Ling, H., et al., *Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase.* Proc. Natl. Acad. Sci. U. S. A., 2004. **101**(8): p. 2265-9.
26. Wang, L. and S. Broyde, *A new anti conformation for N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (AAF-dG) allows Watson-Crick pairing in the Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4).* Nucleic Acids Res., 2006. **34**(3): p. 785-95.
27. Johnson, R.E., et al., *hRAD30 mutations in the variant form of xeroderma pigmentosum.* Science, 1999. **285**(5425): p. 263-5.
28. Masutani, C., et al., *The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta.* Nature, 1999. **399**(6737): p. 700-4.
29. Zhang, Y., et al., *Error-prone lesion bypass by human DNA polymerase eta.* Nucleic Acids Res., 2000. **28**(23): p. 4717-24.
30. Levine, R.L., et al., *Translesion DNA synthesis catalyzed by human pol eta and pol kappa across 1,N6-ethenodeoxyadenosine.* J. Biol. Chem., 2001. **276**(22): p. 18717-21.
31. Haracska, L., S. Prakash, and L. Prakash, *Replication past O(6)-methylguanine by yeast and human DNA polymerase eta.* Mol Cell Biol, 2000. **20**(21): p. 8001-7.
32. Masutani, C., et al., *Mechanisms of accurate translesion synthesis by human polymerase eta.* Embo. J., 2000. **19**(12): p. 3100-9.
33. Zhang, Y., et al., *Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase iota.* Mol. Cell. Biol., 2000. **20**(19): p. 7099-108.
34. Zhang, Y., et al., *Response of human DNA polymerase iota to DNA lesions.* Nucleic Acids Res., 2001. **29**(4): p. 928-35.
35. Haracska, L., et al., *Targeting of human DNA polymerase iota to the replication machinery via interaction with PCNA.* Proc. Natl. Acad. Sci. U. S. A., 2001. **98**(December 4): p. 14256-14261.
36. Johnson, R.E., et al., *Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions.* Nature, 2000. **406**(6799): p. 1015-9.
37. Washington, M.T., et al., *Human DINB1-encoded DNA polymerase kappa is a promiscuous extender of mispaired primer termini.* Proc Natl Acad Sci U S A, 2002. **99**(4): p. 1910-4.
38. Zhang, Y., et al., *Error-free and error-prone lesion bypass by human DNA polymerase kappa in vitro.* Nucleic Acids Res., 2000. **28**(21): p. 4138-46.
39. Nelson, J.R., C.W. Lawrence, and D.C. Hinkle, *Deoxycytidyl transferase activity of yeast REV1 protein.* Nature, 1996. **382**(6593): p. 729-31.
40. Masuda, Y., et al., *Deoxycytidyl transferase activity of the human REV1 protein is closely associated with the conserved polymerase domain.* J Biol Chem, 2001. **276**(18): p. 15051-8.
41. Zhang, Y., et al., *Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion.* Nucleic Acids Res, 2002. **30**(7): p. 1630-8.
42. Fiala, K.A. and Z. Suo, *Sloppy bypass of an abasic lesion catalyzed by a Y-family DNA polymerase.* J. Biol. Chem., 2007. **282**(11): p. 8199-206.
43. *Hepatitis B*. 2009, World Health Organization.
44. *HIV/AIDS Data and Statistics*. 2009, World Health Organization.

45. Stein, D.S. and K.H. Moore, *Phosphorylation of nucleoside analog antiretrovirals: a review for clinicians.* Pharmacotherapy, 2001. **21**(1): p. 11-34.
46. *Approved HBV Drugs*. 2009, Hepatitis B Foundation.
47. *Drugs Used for Treatment of HIV Infection*. 2009, U.S. Food and Drug Administration.
48. Squires, K.E., *An introduction to nucleoside and nucleotide analogues.* Antivir Ther, 2001. **6 Suppl 3**: p. 1-14.
49. Ramadan, K., I. Shevelev, and U. Hubscher, *The DNA-polymerase-X family: controllers of DNA quality?* Nat Rev Mol Cell Biol, 2004. **5**(12): p. 1038-43.
50. Garcia-Diaz, M., et al., *DNA polymerase lambda (Pol lambda), a novel eukaryotic DNA polymerase with a potential role in meiosis.* J. Mol. Biol., 2000. **301**(4): p. 851-67.
51. Fiala, K.A., W. Abdel-Gawad, and Z. Suo, *Pre-Steady-State Kinetic Studies of the Fidelity and Mechanism of Polymerization Catalyzed by Truncated Human DNA Polymerase lambda.* Biochemistry, 2004. **43**(21): p. 6751-62.
52. Ramadan, K., et al., *DNA polymerase lambda from calf thymus preferentially replicates damaged DNA.* J. Biol. Chem., 2002. **277**(21): p. 18454-8.
53. Lee, J.W., et al., *Implication of DNA polymerase lambda in alignment-based gap filling for nonhomologous DNA end joining in human nuclear extracts.* J. Biol. Chem., 2004. **279**(1): p. 805-11.